PREDICTION AND SIGNAL EXTRACTION OF STRONGLY DEPENDENT PROCESSES IN THE FREQUENCY DOMAIN*

by

J Hidalgo
London School of Economics and Political Science

and

Y. Yajima
University of Tokyo

Contents:

The Suntory Centre
Suntory and Toyota International Centres
for Economics and Related Disciplines
London School of Economics and Political Science
Houghton Street
London WC2A 2AE
Tel.: 020-7955 6698

Discussion Paper
No.EM/01/418
June 2001

---

* The first author gratefully acknowledges the research support of the Economic and Social Research Council (ESRC), reference number: R000238212. The second author gratefully acknowledges the research support of the Ministry of Education of Japan, reference number: 09CE20012.

Abstract

We frequently observe that one of the aims of time series analysts is to predict future values of the data. For weakly dependent data, when the model is known up to a finite set of parameters, its statistical properties are well documented and exhaustively examined. However, if the model was misspecified, the predictors would no longer be correct. Motivated by this observation and due to the interest in obtaining adequate and reliable predictors, Bhansali (1974) examined the properties of a nonparametric predictor based on the canonical factorization of the spectral density function given in Whittle (1963) and known as *FLES*.

However, the above work does not cover the so-called strongly dependent data. Due to the interest in this type of process, one of our objectives in this paper is to examine the properties of the *FLES* for these processes. In addition, we illustrate how the *FLES* can be adapted to recover the signal of a strongly dependent process, showing its consistency. The proposed method is semiparametric, in the sense that, in contrast to other methods, we do not need to assume any particular model for the noise except that it is weakly dependent.

Keywords: Prediction; strong dependence; spectral density function; canonical factorization; signal extraction.

JEL Nos.: C22, C53.

Contact address: Dr Javier Hidalgo, Department of Economics, London School of Economics and Political Science, Houghton Street, London WC2 2AE, UK. Email: f.j.hidalgo@lse.ac.uk

1. INTRODUCTION

In empirical studies we often observe that one of the aims to model time series data is for prediction purposes. In the context of linear models, the Wiener-Kolmogorov theory, see for instance Hannan $(1970, Ch.3)$, has been described as a great achievement in that direction. For weakly (linear) dependent data, the statistical properties of predicted future values are very well documented. Existing procedures lie in two main categories, namely the parametric and nonparametric approach. In the former, the parameters of the model are estimated, via either time or frequency domain methods, and plugged into the Wiener-Kolmogorov formula. However, since always there is some degree of uncertainty about the correct specification of the model, Bhansali $(1974, 1977)$ described a nonparametric predictor, based on a factorization of a "windowed" estimate of the spectral density function, which computes the constants which enter in the Wiener-Kolmogorov formula. The algorithm, denoted $FLES$ (factorized logarithm of the estimated spectrum), guarantees the predictor always to be consistent with no need for the practitioner to decide any specific model for the data.

On the other hand, it has been observed that in many areas, such as hydrology or economics, the data may exhibit strong dependence, characterized by having a non-summable autocovariance function. Statistical properties of predictors in parametric models exhibiting strong dependence has not been studied as deeply as with weakly dependent data, although some research has been done in that direction. Among them, we can mention Peiris and Perera (1988), Beran (1994), Crato and Ray (1996). Some empirical examples of prediction with strongly dependent data can be found in Porter-Hudak (1990) and Ray (1993). The former showed the superiority of predictors based on a parametric fractional autoregressive integrated moving average ($FARIMA\,(p, \alpha/2, q)$) model for USA data of monetary aggregates compared to predictors based on more standard autoregressive integrated moving average ($ARIMA$) models.

Because, as was mentioned above, there is always a degree of uncertainty about the correct specification of the data, the first objective of this paper is to study and examine the properties of the predictor based on the $FLES$ algorithm of a covariance stationary linear, possibly, strongly dependent process.

The second objective of the paper is to illustrate how the $FLES$ algorithm can be adapted for the purpose of signal recovering of a covariance stationary linear process which exhibits strong dependence. More specifically, assuming that the process can be decomposed in such a way that the noise is weakly dependent and the signal strongly dependent, we describe how we can extract the signal without assuming any parametric model of the noise. The latter is of interest since one characteristic of many time series, say y_t, is that they can be represented as

$$y_t = p_t + z_t$$

where p_t is the trend, which corresponds to the long run movements, and z_t the irregular (noise) component, which can be regarded as short-term movements. In addition, as many time series are observed quarterly or monthly, it is expected that y_t will have an additional component, say, s_t which represents the seasonal movements of the series, so that

$$y_t = p_t + s_t + z_t.$$

One important issue is the extraction of p_t and/or s_t, see for instance Cleveland and Tiao (1976).

The organization of the paper is as follows. In the next section, we present the $FLES$ algorithm and its estimator. In Section 3, we delimit our framework and examine the properties of the estimators given in Section 2. Section 4 describes how the $FLES$ algorithm can be adapted to extract the signal of a strongly dependent process, showing its statistical properties. Section 5 generalizes the results to cover series with strongly dependent components at other frequencies different than zero, which may be the situation with cyclical/seasonal data. In Section 6, we provide the proofs of our results which apply some technical lemmas given in Section 7. Finally, in the last section, we present conclusions and possible extensions.

2. THE FLES ALGORITHM AND ITS ESTIMATOR

Let $\{x_t\}$ be a covariance stationary linear process which is observed at times $t = 1, ..., n$, having mean that is zero and with absolute continuous spectral distribution, so that its spectral density function, denoted $f_x(\lambda)$, is defined as

$$\gamma_x(j) \overset{def}{=} E(x_0 x_j) = \int_{-\pi}^{\pi} f_x(\lambda)\cos(j\lambda)\, d\lambda, \quad j = 0, 1, 2, ...$$

We will assume that the process x_t admits the following representations

$$x_t = \sum_{j=0}^{\infty} b_j \varepsilon_{t-j}, \ \sum_{j=0}^{\infty} b_j^2 < \infty, \ b_0 = 1 \tag{1}$$

and

$$\sum_{j=0}^{\infty} a_j x_{t-j} = \varepsilon_t, \ a_0 = 1, \tag{2}$$

where ε_t is a martingale difference sequence with mean zero and variance σ_ε^2, and a_j and b_j are constants. Following (1) or (2), the spectral density function can be written as

$$f_x(\lambda) = \frac{\sigma_\varepsilon^2}{2\pi}|A(\lambda)|^{-2} = \frac{\sigma_\varepsilon^2}{2\pi}|B(\lambda)|^2$$

where $A(\lambda) = \sum_{j=0}^{\infty} a_j e^{-ij\lambda}$ is the spectral transfer function of the coefficients a_j. Likewise, $B(\lambda)$ is the transfer function of the coefficients b_j. All throughout the basic assumption of $f_x(\lambda)$ is that

$$f_x(\lambda) \sim C\lambda^{-\alpha} \text{ as } \lambda \to 0+ \tag{3}$$

where $C \in (0, \infty)$, $\alpha \in [0, 1)$ and " $\sim$ " means that the ratio of the left- and right-hand sides tends to one and is differentiable in any open set outside the origin. When $\alpha = 0$, we say that the data is weakly dependent, whereas for $\alpha \in (0, 1)$, we say that the data exhibits the property of strong dependence.

Examples of processes with $\alpha = 0$ are the familiar autoregressive moving average ($ARMA(p, q)$) and Bloomfield's (1973) Exponential processes, whereas examples with $\alpha \in (0, 1)$, we can mention the $FARIMA(p, \alpha/2, q)$ and the Bloomfield's fractional integrated exponential model, see Granger and Joyeux (1980) and Hosking (1981) and Robinson (1994) respectively. Thus, our framework simultaneously allows for both weakly and strongly dependent processes. The latter two models have a spectral density function defined as

$$f_x(\lambda) = \frac{\sigma_\varepsilon^2}{2\pi} \left|1 - e^{i\lambda}\right|^{-\alpha} \left|\frac{\Theta\left(e^{i\lambda}\right)}{\Phi\left(e^{i\lambda}\right)}\right|^2, \; -\pi < \lambda \leq \pi,$$

where $\Phi(\cdot)$ and $\Theta(\cdot)$ are the AR and MA polynomials, respectively, having no zeroes in or on the unit circle, and

$$f_x(\lambda) = \left|1 - e^{i\lambda}\right|^{-\alpha} \exp\left[\sum_{k=1}^{p} \beta_k \cos\{(k-1)\lambda\}\right], \; -\pi < \lambda \leq \pi, \tag{4}$$

respectively.

An earlier example of a process exhibiting the property of strong dependence is the fractional Gaussian noise model introduced by Mandelbrot and Van Ness (1968), whose spectral density function, obtained by Sinai (1976), is

$$f_x(\lambda) = \frac{4\sigma_x^2 \Gamma(\alpha)}{(2\pi)^{3+\alpha}} \cos(\pi\alpha/2) \sin^2(\lambda/2) \sum_{j=-\infty}^{\infty} \left|j + \frac{\lambda}{2\pi}\right|^{-2-\alpha}, \; -\pi < \lambda \leq \pi,$$

where $\sigma_x^2 = Var(x_t)$ and $\Gamma(\cdot)$ is the gamma function. A common feature of all the above models is that their spectral density functions satisfy (3).

Given observations $\{x_{n-j}, j = 1, 2, ...\}$ on the infinite past of the series x_t, let the linear predictor of $x_{n+h}\,(h = 0, 1, ...)$ be denoted by $\widehat{x}_{n+h}$ and the mean-square prediction error by σ_{h+1}^2. Then

$$\widehat{x}_n = -\sum_{u=1}^{\infty} a_u x_{n-u} \quad \text{and} \quad \widehat{x}_{n+h} = -\sum_{u=1}^{h} a_u \widehat{x}_{n+h-u} - \sum_{u=1}^{\infty} a_{u+h} x_{n-u} \tag{5}$$

with

$$\sigma_{h+1}^2 = \sigma_\varepsilon^2 \sum_{u=0}^{h} b_u^2. \tag{6}$$

We notice that as $h \to \infty$ the mean-square prediction error approaches the variance of x_t. So, as $h \to \infty$, knowledge of the past does not help to predict future values.

It is clear that if the coefficients a_j were known, the prediction problem would be solved. Similarly, if $f_x(\lambda)$ was known, these coefficients a_j could be obtained by using the canonical factorization of the spectral density function, see Whittle (1963, p.26) or Brillinger's (1981) Theorem 3.8.4 or Hannan's (1970, p.147) Theorem 6. Specifically, we have that

$$a_j = \frac{1}{2\pi}\int_{-\pi}^{\pi} A(\lambda)\, e^{ij\lambda} d\lambda, \tag{7}$$

$$\sigma_\varepsilon^2 = 2\pi e^{c_0}, \tag{8}$$

where

$$A(\lambda) = \exp\left\{-\sum_{u=1}^{\infty} c_u e^{-iu\lambda}\right\} \tag{9}$$

and

$$c_u = \frac{1}{\pi}\int_0^{\pi} \log\left(f_x(\lambda)\right)\cos(u\lambda)\, d\lambda. \tag{10}$$

The coefficients b_j can likewise be obtained from $B(\lambda) = A^{-1}(\lambda)$.

In practice $f_x(\lambda)$ is unknown, so to compute (10) and therefore equations $(7)-(9)$, $f_x(\lambda)$ needs to be estimated. To that end, introduce the periodogram of x_t

$$I_x(\lambda) = (2\pi n)^{-1}\left|\sum_{t=1}^{n} x_t e^{-it\lambda}\right|^2.$$

We estimate $f_x(\lambda)$ by

$$\widehat{f}_x(\lambda) = \frac{|\lambda|^{-\widehat{\alpha}}}{2m+1}\sum_j |\lambda_j + \lambda|^{\widehat{\alpha}}\, I_x(\lambda + \lambda_j) \tag{11}$$

where $\lambda_j = (2\pi j)/n$, $j = 0, 1, ..., n-1$, $m = m(n)$ a number which increases slowly with n, $\sum_j = \sum_{j=-m}^{m}$, and $\widehat{\alpha}$ is a semiparametric estimator of α, for instance that obtained in Robinson (1995), that is

$$\widehat{\alpha} = \underset{\alpha\in(-1,1)}{\arg\min}\left(\log\left\{\frac{1}{m}\sum_{j=1}^{m}\lambda_j^{\alpha} I_j\right\} - \alpha\frac{1}{m}\sum_{j=1}^{m}\log\lambda_j\right), \tag{12}$$

where $I_j = I(\lambda_j)$.

In the estimator (11) $I_x(\lambda)$ is damped around zero frequency prior to the usual periodogram averaging (which is of the sort stressed by Brillinger, 1981), whereas $\widehat{f}_x$

will typically exhibit a pole at zero frequency. We can regard the estimator in (11) as a prewhitened estimator in the frequency domain, in contrast to that in the time domain suggested in Press and Tukey (1956) when f is believed to have sharp peaks, as is our case. Moreover, see Lemma 1 in Section 7, the estimator (11) possesses better bias properties compared to the usual average periodogram estimate.

Let $\widetilde{\lambda}_j = (\pi j)/M$, $j = 0, \pm 1, ..., \pm M$, where $M = [n/4m]$ and $[z]$ indicates the integer part of z. Abbreviating $\phi\left(\widetilde{\lambda}_\ell\right)$ by ϕ_ℓ, for a generic function $\phi(\lambda)$, (7) $-$ (10) are then estimated by

$$\widehat{c}_u = \frac{1}{M}\sum_{\ell=1}^{M-1}\left(\log \widehat{f}_{x,\ell}\right)\cos\left(u\widetilde{\lambda}_\ell\right), \qquad u = 0, 1, ..., M-1, \tag{13}$$

$$\widehat{A}_j = \exp\left\{-\sum_{u=0}^{M-1}\widehat{c}_u e^{-iu\widetilde{\lambda}_j}\right\} = \overline{\widehat{A}_{-j}}, \quad j = 0, 1, ..., M-1, \tag{14}$$

$$\widehat{a}_u = \frac{1}{2M}\sum_{j=-M+1}^{M}\widehat{A}_j e^{iu\widetilde{\lambda}_j}, \qquad u = 1, ..., M-1, \tag{15}$$

$$\widehat{\sigma}_\varepsilon^2 = 2\pi e^{\widehat{c}_0}, \tag{16}$$

where $\overline{d}$ denotes the conjugate of the complex number d. Thus, (5) is estimated by

$$\widehat{x}_n^* = -\sum_{u=1}^{M-1}\widehat{a}_u\widetilde{x}_{n-u} \text{ and } \widehat{x}_{n+h}^* = -\sum_{u=1}^{h}\widehat{a}_u\widehat{x}_{n+h-u}^* - \sum_{u=1}^{M-h-1}\widehat{a}_{u+h}\widetilde{x}_{n-u} \text{ for } h \geq 1 \tag{17}$$

where $\{\widetilde{x}_{n-j}, j = 1, 2, ..., M-1\}$ is a new independent replicate of observations, with the same statistical properties as x_t, not used in the estimation of the spectral density $f_x(\lambda)$.

We conclude this section mentioning that since we only have a finite record of x_t, $\widehat{A}_j$, $\widehat{a}_u$ and $\widehat{\sigma}_\varepsilon^2$ do not actually estimate the true function/parameters $A\left(\widetilde{\lambda}_j\right)$, a_u and σ_ε^2 respectively, but rather their finite function/parameter versions $A_n\left(\widetilde{\lambda}_j\right)$, $a_{u,n}$ and $\sigma_{\varepsilon,n}^2$. The latter can be obtained by replacing in (14) and (16) $\widehat{c}_u$ by $c_{u,n}$, $u = 0, 1, ..., M-1$, and $\widehat{A}_j$ by $A_n\left(\widetilde{\lambda}_j\right)$ in (15). $c_{u,n}$ is obtained from (13) by replacing $\log\left(\widehat{f}_{x,\ell}\right)$ by $\log(f_{x,\ell})$. Following Bhansali (1974), we refer to the above method of prediction, c.f. (17), as the $FLES$ predictor of the data.

3. ASYMPTOTIC PROPERTIES OF THE ESTIMATORS

Before we establish the asymptotic properties of our estimators given in $(13)-(16)$, and thus the predicted future value of x_t in (17), we introduce the following conditions:

C.1 $f_x(\lambda) = \lambda^{-\alpha} g_x(\lambda)$, $0 < \lambda \leq \pi$, where $0 \leq \alpha < 1$ and $g_x(\lambda)$ is a positive, symmetric around zero and twice continuously differentiable function.

C.2

$$x_t = \sum_{j=0}^{\infty} b_j \varepsilon_{t-j}, \quad \sum_{j=0}^{\infty} b_j^2 < \infty, \quad b_0 = 1,$$

where $E(\varepsilon_t | \mathcal{F}_{t-1}) = 0$, $E(\varepsilon_t^2 | \mathcal{F}_{t-1}) = \sigma_\varepsilon^2$, $E\left(|\varepsilon_t|^\ell | \mathcal{F}_{t-1}\right) = \mu_\ell < \infty$, for $\ell = 3$ and 4, where $\mathcal{F}_t$ the $\sigma-$algebra of events generated by ε_s $s \leq t$, and $cum(\varepsilon_{t_1}, \varepsilon_{t_2}, \varepsilon_{t_3}, \varepsilon_{t_4}) = \kappa_4 \sigma_\varepsilon^4$ if $t_1 = t_2 = t_3 = t_4$ and zero otherwise. In addition, x_t^4 is uniformly integrable.

C.3 $B(\lambda)$ is twice continuously differentiable in any open set outside the origin, and satisfies

$$\frac{\partial}{\partial \lambda} |B(\lambda)| = O\left(\lambda^{-1} |B(\lambda)|\right) \text{ as } \lambda \to 0+.$$

C.4 As $n \to \infty$, $m^4/n^3 + n^2/m^3 \to 0$.

Some comments about Conditions $C.1 - C.4$ are in place. Conditions $C.1$ and $C.3$ are common when analyzing processes which may exhibit strong dependence, see Robinson (1995), Hidalgo (2000*a*) or Hidalgo and Yajima (1999), so their comments apply here. A sufficient condition for the last part of $C.2$ is $\sup_t E|x_t|^{4+\tau} < \infty$ for some $\tau > 0$. This last part of $C.2$ is needed in the proof of Theorems 3 and 4, and in particular to justify that, for example, $E(Y_n) \to E(Y)$ if $Y_n \xrightarrow{d} Y$ and Y_n is uniformly integrable, see Serfling (1980, p. 14). Also, it should be noted that the condition $E(\varepsilon_t | \mathcal{F}_{t-1}) = 0$ is equivalent to the assertion that the best linear predictor is the best predictor, in the least squares sense. This is a natural condition since the final purpose of estimation is often linear prediction, as is our case. Finally, Condition

$C.4$ gives upper and lower bounds on the rate of increase to infinity of the smoothing parameter m.

Theorem 1 *Define* $\widehat{\zeta}_j = \widehat{c}_j - c_{j,n}$ *and* $\widehat{\alpha}$ *as in (12). Assuming C.1-C.4, for any finite collection* $j_1 < j_2 < ... < j_q$,

$$n^{1/2}\left(\widehat{\zeta}_{j_1}, ..., \widehat{\zeta}_{j_q}\right) \xrightarrow{d} N\left(0, \Omega_c\right)$$

where Ω_c *is a diagonal matrix whose* $j-th$ *element is* $(1 + (1+\kappa_4)\,\delta_{j_1})$, *where* $\delta_i = 1$ *if* $i = 0$ *and* $= 0$ *otherwise.*

Theorem 1 indicates that the results obtained in Bhansali (1974) for weakly dependent data, that is $\alpha = 0$, hold the same for strongly dependent data. Thus, Theorem 1 generalizes Bhansali's results to any covariance stationary linear process.

Moreover the results of Theorem 1 have some implications regarding the estimation of the parameters of the Bloomfield exponential model (4). Suppose first that $\alpha = 0$. Then

$$f_x(\lambda) = \exp\left[\sum_{j=1}^{p} \beta_j \cos\{(j-1)\lambda\}\right], \quad -\pi < \lambda \leq \pi,$$

so that it is easy to show that $c_{j,n} - c_j = O(M^{-1})$ and thus that β_j are $n^{1/2}$-consistently estimated by $\widehat{c}_j$ given in (13) choosing $m = 4M = n^{1/2}$, which can be used as a first step in a Whittle-type estimate of β_j, $j = 1, ..., p$, which involves a nonlinear minimization algorithm, say via Newton-Raphson. See also Taniguchi (1987) for an alternative estimator of β_j.

Now suppose that α can be different than zero. In this case we can still obtain simple preliminary estimators of α and β_j to be employed in a Whittle-type estimator as we now illustrate. First the semiparametric estimator of α given in (12) is $m^{1/2}$-consistent, see Robinson (1995), which by C.4 it implies that it is $n^{\tau+1/4}$-consistent for some $\tau > 0$. On the other hand, if

$$f_x(\lambda) = \left|1 - e^{i\lambda}\right|^{-\alpha} \exp\left[\sum_{j=1}^{p} \beta_j \cos\{(j-1)\lambda\}\right], \quad -\pi < \lambda \leq \pi$$

following the arguments of (11), it is obvious that

$$\check{f}_x(\lambda) = \frac{1}{2m+1}\sum_j \left|1-e^{i(\lambda_j+\lambda)}\right|^{\widehat{\alpha}} I_x(\lambda+\lambda_j),$$

is a consistent estimator of $\exp\left[\sum_{j=1}^{p}\beta_j \cos\{(j-1)\lambda\}\right]$. So, if we identify β_j as the c_j in (9), and replace $\widehat{f}_{x,\ell}$ by $\check{f}_{x,\ell}$ in (13), $\widehat{c}_u$ is a $n^{1/2}$-consistent estimator of β_u. Thus $(\widehat{\alpha}, \widehat{c}_j)$ becomes a simple computationally preliminary estimator of (α, β_j).

From Theorem 1 and a simple application of delta methods, we obtain

Corollary 1 *Let* $\sigma^2_{\varepsilon,n} = 2\pi e^{c_{0,n}}$. *Assuming C.1-C.4,*

$$n^{1/2}\left(\widehat{\sigma}^2_{\varepsilon} - \sigma^2_{\varepsilon,n}\right) \xrightarrow{d} N\left(0, \sigma^4_{\varepsilon}(2+\kappa_4)\right).$$

Theorem 2 *Define* $\widehat{\zeta}_j = \widehat{A}_j - A_{j,n}$. *Assuming C.1-C.4, for any finite collection* $j_1, ..., j_q$,

$$n^{1/2} M^{-1/2}\left(\widehat{\zeta}_{j_1}, ..., \widehat{\zeta}_{j_q}\right) \xrightarrow{d} N^c(0, \Omega_A)$$

where $N^c(0, \Omega_A)$ *means a complex normal random variable where the* $\ell j - th$ *element of* Ω_A *is* $2^{-1}\left(\delta_{|j-\ell|} + 2^{-1}\phi_j\phi_\ell - i\phi_{|j-\ell|} + \frac{K^2}{2}\phi_j\phi_\ell\right) A_\ell \overline{A}_j$ *is the* $\ell j - th$ *element of* Ω_A *with*

$\phi_j = (1-\cos(j\pi))/j\pi$ *if* $j > 0$ *and* 1 *if* $j = 0$, *and* $K = \lim_{M\to\infty}\sum_{\ell=1}^{M-1} g_{x,\ell}\left(\int_{-1}^{1}\log\left(1+\frac{v}{2\ell}\right)dv\right)$.

Theorem 3 *Define* $\widehat{\zeta}_j = \widehat{a}_j - a_{j,n}$. *Assuming C.1-C.4, for any finite collection* $j_1, ..., j_q$,

$$n^{1/2}\left(\widehat{\zeta}_{j_1}, ..., \widehat{\zeta}_{j_q}\right) \xrightarrow{d} N(0, \Omega_a)$$

where the $\ell j - th$ *element of* Ω_a *is* $2^{-1}\sum_{p=0}^{\min(j,\ell)} a_{j-p}a_{\ell-p}$.

Once we have obtained the asymptotic properties of the estimators of a_j, $j = 1, ..., M-1$, we are in the position to study the asymptotic properties of the predictor of the data.

Theorem 4 *Assuming C.1-C.4,*

$$\begin{aligned}(a)\;\; AE\left(\widehat{x}^*_n - \widetilde{x}_n\right)^2 &= \sigma^2_{\varepsilon} \\ (b)\;\; AE\left(\widehat{x}^*_{n+h} - \widetilde{x}_{n+h}\right)^2 &= \sigma^2_{h+1} \quad \text{for } h = 1, ..., V \text{ with } V \geq 1\end{aligned}$$

where AE denotes the expectation of the limit distribution.

Theorem 4 illustrates that once again the results obtained by Bhansali (1974) for weakly dependent data extrapolate to data which may exhibit strong dependence.

4. SIGNAL EXTRACTION

4.1. Statement of the problem and parametric estimation of the signal

The problem that we are interested in is as follows. Suppose that a covariance stationary linear process y_t can be decomposed as

$$y_t = x_t + z_t, \quad t = 0, \pm 1, \pm 2, ... \tag{18}$$

where the signal x_t satisfies that

$$(1 - L)^{\alpha/2} x_t = \varepsilon_t^x \tag{19}$$

with $\alpha \in (0, 1)$ and the noise z_t follows the model

$$z_t = \sum_{j=0}^{\infty} a_j^z \varepsilon_{t-j}^z, \ \sum_{j=0}^{\infty} \left| a_j^z \right| < \infty, \ \ a_0^z = 1, \tag{20}$$

where ε_t^x and ε_t^z are white noise mutually independent processes. So, the spectral density function of y_t is

$$f_y (\lambda) = f_x (\lambda) + f_z (\lambda) \tag{21}$$

where $f_w (\lambda)$ denotes the spectral density function of a generic covariance stationary linear process w_t. As an example, suppose that (20) follows an $ARMA\,(p, q)$ process. Then

$$f_y (\lambda) = \frac{1}{2\pi} \left(\sigma_{\varepsilon^x}^2 \left| 1 - e^{i\lambda} \right|^{-\alpha} + \sigma_{\varepsilon^z}^2 \left| \frac{\Theta \left(e^{i\lambda} \right)}{\Phi \left(e^{i\lambda} \right)} \right|^2 \right),$$

where $\Phi (L)$ and $\Theta (L)$ are the AR and MA polynomials, respectively, with no common roots and having no zeroes in or on the unit circle.

The purpose of this section is, given the observed data y_t, $t = 1, .., n$, to estimate the signal x_t. It is known that if all the past and future values of y_t were observed, then the Kolmogorov-Wiener formula would provide the best linear predictor (BLP) of x_t given y_s, $s = 0, \pm 1, \pm 2, ...$, defined by

$$x_{t|\infty} = \sum_{j=-\infty}^{\infty} \psi_j y_{t-j}$$

where ψ_j minimizes $E\left(x_t - \sum_{j=-\infty}^{\infty} \psi_j y_{t-j}\right)^2$ and equals to the jth Fourier coefficient of $f_y^{-1}(\lambda) f_x(\lambda)$, that is

$$\psi_j = \frac{1}{2\pi} \int_{-\pi}^{\pi} f_y^{-1}(\lambda) f_x(\lambda) e^{ij\lambda} d\lambda.$$

However, in empirical studies, because not all values of y_t are observed, some truncation will be needed to implement $x_{t|\infty}$. We shall denote the estimate of $x_{t|\infty}$ by

$$\widehat{x}_{t|\infty} = \widehat{E\left[x_t \left| y_1, ..., y_n\right.\right]} = \sum_{j=t-n}^{t-1} \widehat{\psi}_j y_{t-j} \tag{22}$$

where $\widehat{\psi}_j$ is some estimate of ψ_j.

When a full parameterization of the process generating z_t is known, for example z_t follows an $ARMA(p, q)$, where $\Theta(L) = \Theta(L; \vartheta)$ and $\Phi(L) = \Phi(L; \vartheta)$, then $f_y(\lambda) = f_y(\lambda; \tau)$, for all $\lambda \in (0, \pi]$, where $\tau = (\alpha, \sigma_{\varepsilon^x}^2, \sigma_{\varepsilon^z}^2, \vartheta')'$. If the parameters τ were known, the signal extraction problem would be solved by plugging those values of τ into the right side of (22) to obtain

$$\widehat{x}_{t|\infty}(\tau) = \sum_{j=t-n}^{t-1} \psi_j(\tau) y_{t-j}. \tag{23}$$

However, in practice τ is unknown, and thus to implement (23), τ is replaced by, for example, the Whittle estimate $\widetilde{\tau}$ which, under suitable conditions, is known to be $n^{1/2}$-consistent and asymptotically normal, see Fox and Taqqu (1986), Dahlhaus (1989), Giraitis and Surgailis (1990) or Hosoya (1997). Given $\widetilde{\tau}$, $x_{t|\infty}$ is then estimated by plugging $\widetilde{\tau}$ into the right side of (23), obtaining

$$\widehat{x}_{t|\infty}(\widetilde{\tau}) = \sum_{j=t-n}^{t-1} \psi_j(\widetilde{\tau}) y_{t-j}.$$

The parametric approach suffers from two possible drawbacks. First, its implementation can be difficult since to obtain the "reduced" form parameters of $f_y(\lambda; \tau)$ can be quite complex, even in simpler situations than those considered here. The second, and possibly more important, drawback is that the procedure described is very sensitive to a correct specification of the spectral density function of y_t, that is $f_y(\lambda)$. In particular, on a correct specification of the order of the polynomials $\Phi(L)$ and $\Theta(L)$ if indeed the noise z_t followed an $ARMA$ process. For instance, z_t might follow a Bloomfield's exponential model instead of an $ARMA$ one. If that was the case it would lead to inconsistent estimates of $f_y(\lambda)$ and so, the estimates of $x_{t|\infty}$, $\widehat{x}_{t|\infty}$, would be inadequate and "inconsistent".

Looking at equations (18), (19) and (20), we can regard our setup as semiparametric, in that only the term which we are interested in is parameterized. That is, we have a parametric model for the underlying structure of the signal x_t, while we have left the model for the noise z_t unspecified. So, in the terminology employed in semiparametric statistics, we can consider the parameters $(\sigma^2_{\varepsilon^z}, \vartheta')'$ as nuisance parameters. Thus, the question of interest is whether we can "estimate" the parametric part of the model, that is, to extract the signal x_t, in the presence of those nuisance parameters represented by the noise process z_t. This is answered in the next subsection.

4.2. Semiparametric estimation of the signal x_t

Our main concern lies in the estimation of $x_{t|\infty}$, that is $\widehat{x}_{t|\infty}$. Assuming that y_t follows (18), (19) and (20),

$$\psi_j = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\sigma^2_{\varepsilon^x} \left|1 - e^{i\lambda}\right|^{-\alpha}}{\sigma^2_{\varepsilon^x} \left|1 - e^{i\lambda}\right|^{-\alpha} + (2\pi) f_z(\lambda)} e^{ij\lambda} d\lambda = \frac{\sigma^2_{\varepsilon^x}}{2\pi} \int_{-\pi}^{\pi} g^{-1}(\lambda) e^{ij\lambda} d\lambda,$$

where $g(\lambda) = \sigma^2_{\varepsilon^x} + (2\pi) \left|1 - e^{i\lambda}\right|^{\alpha} f_z(\lambda)$ times $(2\pi)^{-1}$ can be regarded as the spectral density function of $(1-L)^{\alpha/2} y_t$. Assuming that y_t is a covariance stationary linear process which admits a representation as an infinite autoregressive model as in (2),

$(1-L)^{\alpha/2} y_t = w_t$ will have also such a representation, say

$$\sum_{j=0}^{\infty} \beta_j w_{t-j} = \varepsilon_t^w; \quad \beta_0 = 1.$$

Following the results of Section 3, $g\left(\widetilde{\lambda}_p\right)$ can be estimated by

$$\widehat{g}\left(\widetilde{\lambda}_p\right) = \widehat{\sigma}^2_{\varepsilon^w} \left|1 + \sum_{j=1}^{N} \widehat{\beta}_j e^{ij\widetilde{\lambda}_p}\right|^{-2}$$

where $\widehat{\sigma}^2_{\varepsilon^w}$, $\widehat{\beta}_j$, $j = 1, ..., N$, are obtained by means of the $FLES$ algorithm described in $(13) - (16)$ with $f_x(\lambda)$ being replaced by $g(\lambda)$. So ψ_j can be estimated by

$$\widehat{\psi}_j = \frac{\widehat{\sigma}^2_{\varepsilon^x}}{2M} \sum_{p=-M}^{M-1} \widehat{g}^{-1}\left(\widetilde{\lambda}_p\right) e^{ij\widetilde{\lambda}_p}, \tag{24}$$

with $\widehat{g}(\lambda)$ an estimate of $g(\lambda)$ and thus $x_{t|\infty}$ is estimated by

$$\widehat{x}_{t|\infty} = \sum_{j=\max(t-n,1-N)}^{\min(t-1,N-1)} \widehat{\psi}_j y_{t-j}. \tag{25}$$

Thus, the problem to obtain $\widehat{x}_{t|\infty}$ in (25) is reduced to obtaining estimates of σ^2_ε and α. Given our model (19), $f_x(\lambda) = (\sigma^2_{\varepsilon^x}/2\pi)\left|1 - e^{i\lambda}\right|^{-\alpha}$ which together with (18) and (20), implies that $f_y(\lambda) \sim C\lambda^{-\alpha}$ as $\lambda \to 0+$, c.f. (3). To observe the latter claim, assume that z_t follows an $ARMA$ process for expositional simplicity. From (21), we have that

$$\begin{aligned} f_y(\lambda) &= \frac{1}{2\pi}\left(\sigma^2_{\varepsilon^x}\left|1 - e^{i\lambda}\right|^{-\alpha} + \sigma^2_{\varepsilon^z}\left|\frac{\Theta\left(e^{i\lambda}\right)}{\Phi\left(e^{i\lambda}\right)}\right|^2\right) \\ &\sim \frac{\sigma^2_{\varepsilon^x}}{2\pi}\lambda^{-\alpha} + O(1) \qquad \text{as } \lambda \to 0+ \end{aligned}$$

because $\left|\Theta\left(e^{i\lambda}\right)/\Phi\left(e^{i\lambda}\right)\right|^2$ is continuous for all $\lambda \in [0, \pi]$, and

$$\left|1 - e^{i\lambda}\right|^{-\alpha} = 2^{-\alpha}\left(\sin|\lambda/2|\right)^{-\alpha} \sim \lambda^{-\alpha}\left(1 + C_1\lambda^2\right) \qquad \text{as } \lambda \to 0+,$$

where C_1 is a finite positive constant. Thus, we conclude that

$$f_y(\lambda) \sim \frac{\sigma^2_{\varepsilon^x}}{2\pi}\lambda^{-\alpha} \qquad \text{as } \lambda \to 0+, \tag{26}$$

so that α can be estimated as in Section 2, i.e. employing Robinson's (1995) estimator given in (12), and $\sigma^2_{\varepsilon^x}/2\pi$ by $m^{-1}\sum_{j=1}^{m}\lambda_j^{\widehat{\alpha}} I_j$.

Let us introduce the following condition,

C.5 $N^{-1}+N^{3+\alpha}m^{-2}\rightarrow 0$ and $NM^{-1}\leq 1$.

Observe that from C.4, we can choose N to be equal to M in C.5. However, we leave C.5 in its present form to give somehow more generality to the result in Theorem 5 below.

Theorem 5 *Assuming C.1-C.5, as* $n\rightarrow\infty$, $\widehat{x}_{[n\delta]|\infty}-x_{[n\delta]|\infty}\overset{P}{\rightarrow}0$, *where* $\delta\in(0,1)$.

Theorem 5 indicates that the simply implemented signal extraction algorithm is consistent. However, and more importantly, in the process of performing the signal extraction of x_t, there has been no need to specify any particular structure for the noise z_t. Thus, we have avoided the problem that a bad specification for the noise may induce on the extraction of the signal. We can therefore consider the approach as semiparametric, where z_t, following the semiparametric terminology, is the nuisance "parameter" or function.

5. EXTENSION TO SEASONAL/CYCLICAL DATA

This section generalizes the results obtained in Sections 3 and 4 to data exhibiting cyclical behaviour. This is motivated by the often periodic behaviour exhibited in many time series and which is manifested by a sharp peak on the spectral density function estimate. A model capable to generate strong cyclical/periodic dependence is

$$\left(1-2\left(\cos\lambda^0\right)L+L^2\right)^{d_0}x_t=\varepsilon_t, \tag{27}$$

introduced by Andel (1986) and Gray et al. (1989), where ε_t follows an $ARMA\,(p,q)$ process and $d_0=\alpha/2$ if $\lambda^0\neq 0,\pi$ or $\alpha/4$ if $\lambda^0=0,\pi$. Gray et al. (1989) coined (27) the $GARMA$ process and is characterized by having a spectral density function defined as

$$f_x(\lambda) = \frac{\sigma^2}{2\pi} \left|1 - 2\left(\cos\lambda^0\right)e^{i\lambda} + e^{i2\lambda}\right|^{-2d_0} \left|\frac{\Theta\left(e^{i\lambda};\theta\right)}{\Phi\left(e^{i\lambda};\theta\right)}\right|^2 \qquad -\pi < \lambda \leq \pi \tag{28}$$

where $\Phi(\cdot)$ and $\Theta(\cdot)$ are the AR and MA polynomials, respectively, having no zeroes in or on the unit circle. Similarly, we can generalize (27) allowing ε_t to follow a Bloomfield exponential instead of an $ARMA$ process, having a spectral density function given by

$$f_x(\lambda) = \left|1 - 2\left(\cos\lambda^0\right)e^{i\lambda} + e^{i2\lambda}\right|^{-2d_0} \exp\left[\sum_{k=1}^{p} \beta_k \cos\{(k-1)\lambda\}\right], \; -\pi < \lambda \leq \pi. \tag{29}$$

Models (28) and (29) can be extended to allow for more than one spectral singularity, e.g.

$$f_x(\lambda) = \frac{\sigma^2}{2\pi} \left(\prod_{j=1}^{h} \left|1 - 2\left(\cos\lambda^j\right)e^{i\lambda} + e^{i2\lambda}\right|^{-2d_j}\right) \left|\frac{\Theta\left(e^{i\lambda};\theta\right)}{\Phi\left(e^{i\lambda};\theta\right)}\right|^2 \qquad -\pi < \lambda \leq \pi \tag{30}$$

and

$$f_x(\lambda) = \left(\prod_{j=1}^{h} \left|1 - 2\left(\cos\lambda^j\right)e^{i\lambda} + e^{i2\lambda}\right|^{-2d_j}\right) \exp\left[\sum_{k=1}^{p} \beta_k \cos\{(k-1)\lambda\}\right], -\pi < \lambda \leq \pi,$$

respectively, which may be reasonable for seasonal data (e.g. for monthly data take $h = 7$ and $\lambda^j = (j-1)\pi/6, \; j = 1, ..., 7$). Observe that, for example, for monthly data we have permitted the degree of strong dependence, that is d_j, to differ across the seasonal frequencies. This generalizes the model study by Carlin and Dempster (1989) or Porter-Hudak (1990), who assumed that $d_j = d$ for all $j = 2, ..., 6$ and $d_j = d/2$ for $j = 1, 7$, since for the latter in the time domain (30) becomes

$$\Phi(L,\theta)\left(1 - L^{12}\right)^d x_t = \Theta(L,\theta)\, u_t$$

where u_t is a white noise process.

We now describe the $FLES$ algorithm of Section 2. To that end, and for the sake of presentation, take the number of possible spectral singularities to be one, say at λ^0. Our basic assumption on $f_x(\lambda)$ given in (3) becomes

$$f_x(\lambda) \sim C\left|\lambda - \lambda^0\right|^{-\alpha} \text{ as } \lambda \to \lambda^0 \tag{31}$$

where $C \in (0, \infty)$ and $\alpha \in [0, 1)$. We now estimate $f_x(\lambda)$ by

$$\widehat{f}_x(\lambda) = \frac{\left|\lambda - \lambda^0\right|^{-\widehat{\alpha}}}{2m+1} \sum_j \left|\lambda_j + \lambda - \lambda^0\right|^{\widehat{\alpha}} I_x(\lambda + \lambda_j) \tag{32}$$

where α is estimated as in Arteche and Robinson (2000), that is

$$\widehat{\alpha} = \underset{\alpha \in (-1,1)}{\arg\min} \left(\log \left\{ \frac{1}{2m} \sum_{j=-m, \neq 0}^{m} \left|\lambda_j - \lambda^0\right|^{\alpha} I_x(\lambda_j) \right\} - \frac{\alpha}{2m} \sum_{j=-m, \neq 0}^{m} \log \left|\lambda_j - \lambda^0\right| \right). \tag{33}$$

Given (32) and (33), the $FLES$ algorithm to predict future values of x_t is exactly the same as that in Section 2. That is,

$$\begin{aligned}
\widehat{c}_u &= \frac{1}{M} \sum_{\ell=1}^{M-1} \left(\log \widehat{f}_{x,\ell}\right) \cos\left(u\widetilde{\lambda}_\ell\right), & u = 0, 1, ..., M-1, \\
\widehat{A}_j &= \exp\left\{ -\sum_{u=0}^{M-1} \widehat{c}_u e^{-iu\widetilde{\lambda}_j} \right\} = \overline{\widehat{A}_{-j}}, \quad j = 0, 1, ..., M-1, \\
\widehat{a}_u &= \frac{1}{2M} \sum_{j=-M+1}^{M} \widehat{A}_j e^{iu\widetilde{\lambda}_j}, & u = 1, ..., M-1, \\
\widehat{\sigma}_\varepsilon^2 &= 2\pi e^{\widehat{c}_0},
\end{aligned} \tag{34}$$

where $\widehat{f}_{x,\ell} = \widehat{f}_x(\lambda_\ell)$ given in (32). Then, (5) is estimated by

$$\widehat{x}_n^* = -\sum_{u=1}^{M-1} \widehat{a}_u \widetilde{x}_{n-u} \text{ and } \widehat{x}_{n+h}^* = -\sum_{u=1}^{h} \widehat{a}_u \widehat{x}_{n+h-u}^* - \sum_{u=1}^{M-h-1} \widehat{a}_{u+h} \widetilde{x}_{n-u} \text{ for } h \geq 1$$

where $\{\widetilde{x}_{n-j}, j = 1, 2, ..., M-1\}$ is a new independent replicate of observations, with the same statistical properties as x_t, not used in the estimation of the spectral density $f_x(\lambda)$.

Let us introduce the following conditions:

C.1' $f_x(\lambda) = \left|\lambda - \lambda^0\right|^{-\alpha} g_x(\lambda)$, $0 < \lambda \leq \pi$, where $0 \leq \alpha < 1$ and $g_x(\lambda)$ is a positive, symmetric around zero and twice continuously differentiable function.

C.3' $B(\lambda)$ is twice continuously differentiable in any open set outside λ^0 and satisfies

$$\frac{\partial}{\partial \lambda} |B(\lambda)| = O\left(\left|\lambda - \lambda^0\right|^{-1} |B(\lambda)|\right) \text{ as } \lambda \to \lambda^0.$$

Theorem 6 *Define* $\widehat{\zeta}_j = \widehat{c}_j - c_{j,n}$ *and* $\widehat{\alpha}$ *as in* (33)*. Assuming C.1',C.2, C.3' and C.4, for any finite collection* $j_1 < j_2 < ... < j_q$,

$$n^{1/2}\left(\widehat{\zeta}_{j_1}, ..., \widehat{\zeta}_{j_q}\right) \xrightarrow{d} N\left(0, \Omega_c\right)$$

where Ω_c *is a diagonal matrix whose* $j-th$ *element is* $\left(1 + \left(1 + \kappa_4\right) \delta_{j_1}\right)$.

Proof. The proof of this result or any other in this section follows proceeding as with the proofs of those results in Sections 3 and 4 and so, they are omitted. □

Theorem 7 *Assuming C.1',C.2, C.3' and C.4,*

$$\begin{aligned} (a)\ AE\left(\widehat{x}_n^* - \widetilde{x}_n\right)^2 &= \sigma_\varepsilon^2 \\ (b)\ AE\left(\widehat{x}_{n+h}^* - \widetilde{x}_{n+h}\right)^2 &= \sigma_{h+1}^2 \quad \text{for } h = 1, ..., V \text{ with } V \geq 1. \end{aligned}$$

We now turn our attention to the signal extraction problem. Suppose that a covariance stationary linear process y_t, which is observed at times $t = 1, ..., n$, is decomposed as

$$y_t = x_t + z_t, \quad t = 0, \pm 1, \pm 2, ... \tag{35}$$

where x_t and z_t denote the signal and noise respectively. Assume that x_t and z_t satisfy

$$g\left(L, \alpha\right) x_t = \varepsilon_t^x \tag{36}$$

denoting $g\left(L, \alpha\right) = \left(1 - 2\left(\cos \lambda^0\right) L + L^2\right)^{\alpha/2}$ with $\alpha \in (0, 1)$ where

$$z_t = \sum_{j=0}^{\infty} a_j^z \varepsilon_{t-j}^z, \ \sum_{j=0}^{\infty} \left|a_j^z\right| < \infty, \ a_0^z = 1, \tag{37}$$

respectively, where ε_t^x and ε_t^z are white noise mutually independent processes. So, the spectral density function of y_t is $f_y\left(\lambda\right) = f_x\left(\lambda\right) + f_z\left(\lambda\right)$. As an example, suppose that (37) follows an $ARMA\left(p, q\right)$ process, then

$$f_y\left(\lambda\right) = \frac{1}{2\pi}\left(\sigma_{\varepsilon^x}^2 \left|1 - 2\left(\cos \lambda^0\right) e^{i\lambda} + e^{2i\lambda}\right|^{-\alpha} + \sigma_{\varepsilon^z}^2 \left|\frac{\Theta\left(e^{i\lambda}\right)}{\Phi\left(e^{i\lambda}\right)}\right|^2\right).$$

The objective is, given the observed data y_t, $t = 1, .., n$, to estimate the signal x_t. In this case, ψ_j is defined as

$$\psi_j = \frac{\sigma^2_{\varepsilon^x}}{2\pi} \int_{-\pi}^{\pi} g^{-1}(\lambda)\, d\lambda.$$

where $g(\lambda) = (2\pi) \left|1 - 2\left(\cos \lambda^0\right) e^{i\lambda} + e^{2i\lambda}\right|^{\alpha} f_y(\lambda)$.

Using the convention that $y_t = 0$ for $t \leq 0$ or $t > n$, the results of Section 3 and 4 suggest to estimate $x_{t|\infty}$ by

$$\widehat{x}_{t|\infty} = \sum_{j=\max(t-n,1-N)}^{\min(t-1,N-1)} \widehat{\psi}_j y_{t-j} \tag{38}$$

where

$$\widehat{\psi}_j = \frac{\widehat{\sigma}^2_{\varepsilon^x}}{2M} \sum_{p=-M}^{M-1} \widehat{g}^{-1}\left(\widetilde{\lambda}_p\right) e^{ij\widetilde{\lambda}_p}.$$

Assuming that y_t is a covariance stationary linear process which admits a representation as an infinite autoregressive model as in (2), $w_t = \left(1 - 2\left(\cos \lambda^0\right) L + L^2\right)^{\alpha/2} y_t$ will have also such a representation, say $\sum_{j=0}^{\infty} \beta_j w_{t-j} = \varepsilon^w_t$; and $\beta_0 = 1$.

Following the results of Section 3, $g\left(\widetilde{\lambda}_p\right)$ can be estimated by

$$\widehat{g}\left(\widetilde{\lambda}_p\right) = \widehat{\sigma}^2_{\varepsilon^w} \left|1 + \sum_{j=1}^{N} \widehat{\beta}_j e^{ij\widetilde{\lambda}_p}\right|^{-2}$$

where $\widehat{\sigma}^2_{\varepsilon^w}$, $\widehat{\beta}_j$, $j = 1, ..., N$, are estimated using the $FLES$ algorithm described in $(13) - (16)$ where $f_x(\lambda)$ is replaced by $g(\lambda)$. Thus, the problem to obtain $\widehat{x}_{t|\infty}$ in (38) is reduced to obtaining estimates of $\sigma^2_{\varepsilon^x}$ and α.

Indeed, given our model (36), $f_x(\lambda) = \left(\sigma^2_{\varepsilon^x}/2\pi\right) \left|1 - 2\left(\cos \lambda^0\right) e^{i\lambda} + e^{2i\lambda}\right|^{-\alpha}$ which together with (35) and (37), implies that $f_y(\lambda) \sim C\left|\lambda - \lambda^0\right|^{-\alpha}$ as $\lambda \to \lambda^0$, by similar algebra to that in Section 4.2. Thus, we conclude that

$$f_y(\lambda) \sim \frac{\sigma^2_{\varepsilon^x}}{2\pi} \left|\lambda - \lambda^0\right|^{-\alpha} \qquad \text{as } \lambda \to \lambda^0, \tag{39}$$

so α can be estimated proceeding as in Section 4.2 but employing Arteche and Robinson's (2000) estimator given in (33) and $\sigma^2_{\varepsilon^x}/2\pi$ by $(2m)^{-1} \sum_{j=-m,\neq 0}^{m} \lambda_j^{\widehat{\alpha}} I\left(\lambda^0 + \lambda_j\right)$.

Theorem 8 *Assuming C.1', C.2, C.3', C.4 and C.5, as* $n \to \infty$, $\widehat{x}_{[n\delta]|\infty} - x_{[n\delta]|\infty} \xrightarrow{P} 0$, *where* $\delta \in (0,1)$.

We finish this section with two remarks.

Remark 1 *If* $x_t = p_t + s_t$ *where* p_t *and* s_t *follow the models* (19) *and* (36) *respectively, then*

$$f_x(\lambda) = \frac{\sigma_1^2}{2\pi} \left|1 - e^{i\lambda}\right|^{-\alpha_1} + \frac{\sigma_2^2}{2\pi} \left|1 - 2\left(\cos\lambda^0\right) e^{i\lambda} + e^{i2\lambda}\right|^{-\alpha_2}.$$

If this were the case α_1 *and* α_2 *can still be estimated by* (12) *and* (33) *respectively. The reason is because the estimators only involve frequencies in a shrinking neighbourhood of* 0 *and* λ^0 *respectively, and* $\left|\lambda^0\right| > \delta$ *for some* $\delta > 0$.

Remark 2 *The results/algorithm discussed in this section assume that, say* λ^0, *is known. However, with real data that knowledge can be no so obvious, as when investigating the length of a cycle in macroeconomic time series. In this case, we can appeal to results in Yajima* (1996) *or Hidalgo* (2000*b*). *In particular, Hidalgo* (2000*b*) *provides a semiparametric estimator of* λ^0 *which is* $n^{\delta+1/2}-$*consistent for some* $\delta \in (0, 1/2)$ *and asymptotically normal. Moreover, Hidalgo* (2000*b*) *shows that the properties of the semiparametric estimator of* α_1 *and/or* α_2 *are the same as when* λ^0 *is known.*

6. PROOFS

Proof of Theorem 1 By Wold device, it suffices to show that for finite constants φ_j,

$$n^{1/2} \sum_{j=p}^{q} \varphi_j \left(\widehat{c}_j - c_{j,n}\right) \xrightarrow{d} N\left(0, \sum_{j=p}^{q} \left(1 + (1+\kappa_4)\,\delta_j\right) \varphi_j^2\right). \tag{40}$$

Using the definitions of $c_{j,n}$ and $\widehat{c}_j$, a typical component on the left of (40) is

$$\left(\widehat{c}_j - c_{j,n}\right) = \frac{1}{M} \sum_{\ell=1}^{M-1} \left(\log \widehat{f}_{x,\ell} - \log \widetilde{f}_{x,\ell}\right) \cos\left(j\widetilde{\lambda}_\ell\right) + \left(\widetilde{c}_{j,n} - c_{j,n}\right), \tag{41}$$

where, for $j = 0, 1, ..., M-1$, $\widetilde{c}_{j,n} = M^{-1}\sum_{\ell=1}^{M-1}\left(\log \widetilde{f}_{x,\ell}\right)\cos\left(j\widetilde{\lambda}_{\ell}\right)$ with $\widetilde{f}_{x,\ell} = \lambda_{2m\ell}^{-\alpha}(2m+1)^{-1}\sum_{p} g_x\left(\lambda_{p+2m\ell}\right)$.

First by Lemma 1 and $C.4$ the second term on the right of (41) is $o\left(n^{-1/2}\right)$. Next, since $\sup_{\ell} a_{\ell}^2 \leq \sum_{\ell} a_{\ell}^2$, $\sup_{\ell=1,..M}\left|\left(\widehat{f}_{x,\ell} - \widetilde{f}_{x,\ell}\right)/\widetilde{f}_{x,\ell}\right|^2$ is bounded by

$$\sum_{\ell=1}^{M}\left|\left(\widehat{f}_{x,\ell} - \widetilde{f}_{x,\ell}\right)/\widetilde{f}_{x,\ell}\right|^2 \leq 2\sum_{\ell=1}^{M}\left|\left(\widehat{f}_{x,\ell} - \check{f}_{x,\ell}\right)/\widetilde{f}_{x,\ell}\right|^2 + 2\sum_{\ell=1}^{M}\left|\left(\check{f}_{x,\ell} - \widetilde{f}_{x,\ell}\right)/\widetilde{f}_{x,\ell}\right|^2$$

where $\check{f}_{x,\ell} = \lambda_{2m\ell}^{-\alpha}(2m+1)^{-1}\sum_{p}\lambda_{j+2m\ell}^{\alpha} I_x\left(\lambda_{p+2m\ell}\right)$. Because $C.4$ and Lemmas 2 and 3 respectively imply that the right side of the last displayed inequality is $o_p(1)$, by Taylor expansion of $\log(x)$ the first term on the right of (41) is

$$\frac{1}{M}\sum_{\ell=1}^{M-1}\left(\frac{\check{f}_{x,\ell} - \widetilde{f}_{x,\ell}}{\widetilde{f}_{x,\ell}}\right)\cos\left(j\widetilde{\lambda}_{\ell}\right) + \frac{1}{2M}\sum_{\ell=1}^{M-1}\left(\frac{\check{f}_{x,\ell} - \widetilde{f}_{x,\ell}}{\widetilde{f}_{x,\ell}}\right)^2\cos\left(j\widetilde{\lambda}_{\ell}\right)(1 + o_p(1))$$
$$+\frac{1}{M}\sum_{\ell=1}^{M-1}\left(\frac{\widehat{f}_{x,\ell} - \check{f}_{x,\ell}}{\widetilde{f}_{x,\ell}}\right)\cos\left(j\widetilde{\lambda}_{\ell}\right) + \frac{1}{2M}\sum_{\ell=1}^{M-1}\left(\frac{\widehat{f}_{x,\ell} - \check{f}_{x,\ell}}{\widetilde{f}_{x,\ell}}\right)^2\cos\left(j\widetilde{\lambda}_{\ell}\right)(1 + o_p(1)), \tag{42}$$

where the $o_p(1)$ is uniformly in ℓ. The second term of (42) is $O_p\left(m^{-1}\right)$ by Lemma 2, whereas the third and fourth terms of (42) are $O_p\left(m^{-1/2}M^{-1} + m^{-1}\right)$ by Lemma 3 and $\left|\sum_{\ell=1}^{M} g_{x,\ell}\int_{-1}^{1}\log\left(1 + \frac{v}{2\ell}\right)dv\right| < \infty$ after straightforward calculations. Thus, we conclude that the first term on the right of (41), i.e. $\widehat{c}_j - \widetilde{c}_{j,n}$, is

$$\frac{1}{M}\sum_{\ell=1}^{M-1}\left(\frac{\check{f}_{x,\ell} - \widetilde{f}_{x,\ell}}{\widetilde{f}_{x,\ell}}\right)\cos\left(j\widetilde{\lambda}_{\ell}\right) + O_p\left(m^{-1/2}M^{-1} + m^{-1}\right),$$

whose first term is

$$\frac{1}{M}\sum_{\ell=1}^{M-1}\left(\frac{\check{f}_{x,\ell} - \widetilde{f}_{x,\ell}}{f_{x,\ell}}\right)\cos\left(j\widetilde{\lambda}_{\ell}\right) + \frac{1}{M}\sum_{\ell=1}^{M-1}\left(\frac{\check{f}_{x,\ell} - \widetilde{f}_{x,\ell}}{f_{x,\ell}}\right)\left(\frac{f_{x,\ell} - \widetilde{f}_{x,\ell}}{\widetilde{f}_{x,\ell}}\right)\cos\left(j\widetilde{\lambda}_{\ell}\right). \tag{43}$$

The first absolute moment of the second term of (43) is bounded by

$$\frac{1}{M}\sum_{\ell=1}^{M-1}\left|\frac{f_{x,\ell} - \widetilde{f}_{x,\ell}}{f_{x,\ell}}\right| E\left|\frac{\check{f}_{x,\ell} - \widetilde{f}_{x,\ell}}{\widetilde{f}_{x,\ell}}\right| = O\left(\frac{1}{M^2 m^{1/2}}\right) = o\left(n^{-1/2}\right)$$

by Lemma 2 and the proof of Lemma 1.

Thus, except negligible ending effects, by definition of $\widehat{f}_{x,\ell}$, $\widetilde{f}_{x,\ell}$ and $f_{x,\ell}$ and $C.4$

$$n^{1/2}\left(\widehat{c}_j - c_{j,n}\right) = \frac{n^{1/2}}{[n/2]} \sum_{s=1}^{[n/2]} \left(\frac{I_x\left(\lambda_s\right) - f_x\left(\lambda_s\right)}{f_x\left(\lambda_s\right)} \right) h_{j,n}\left(s\right) + o_p\left(1\right), \qquad (44)$$

where $h_{j,n}(s) = g_{x,\ell}^{-1} g_x(\lambda_s) \cos\left(j\widetilde{\lambda}_\ell\right)$ if $(2\ell-1)\, m < s \le (2\ell+1)\, m$ and $\ell = 1, ..., M-1$. That is, $h_{j,n}(s)$ is a step function in $[0,\pi]$. Now, by Lemmas 3.1 and 3.2 of Giraitis et al. (2000), which are a simple extension of the proof of (4.8) in Robinson (1995), after observing that $|h_{j,n}(s)|$ is a bounded function,

$$\sum_{s=1}^{[n/2]} \left(\frac{I_x\left(\lambda_s\right)}{f_x\left(\lambda_s\right)} - \frac{(2\pi)\, I_\varepsilon\left(\lambda_s\right)}{\sigma_\varepsilon^2} \right) h_{j,n}\left(s\right) = O_p\left(n^{1/3}\log^{2/3} n\right) \qquad (45)$$

where $I_\varepsilon(\lambda_s)$ denotes the periodogram of the innovations ε_t in (1). So, the right side of (44) is

$$\frac{n^{1/2}}{[n/2]} \sum_{s=1}^{[n/2]} \left(\frac{(2\pi)\, I_\varepsilon\left(\lambda_s\right)}{\sigma_\varepsilon^2} - 1 \right) h_{j,n}\left(s\right) + o_p\left(1\right),$$

and we conclude that the left side of (40) is

$$n^{1/2} \sum_{j=p}^{q} \varphi_j \left(\widehat{c}_j - c_{j,n}\right) = \frac{n^{1/2}}{[n/2]} \sum_{s=1}^{[n/2]} \left(\frac{(2\pi)\, I_\varepsilon\left(\lambda_s\right)}{\sigma_\varepsilon^2} - 1 \right) \sum_{j=p}^{q} \varphi_j h_{j,n}\left(s\right) + o_p\left(1\right).$$

But by an extension of Robinson's (1995) Theorem 2, see Giraitis et al. (2000), the right side converges in distribution to $N(0,V)$ where V is, by the definition of $h_{j,n}(s)$, $s = 1, ..., [n/2]$,

$$\begin{aligned}
\sum_{j_1,j_2=p}^{q} \varphi_{j_1}\varphi_{j_2}\Omega_{c,j_1,j_2} &= \lim_{n\to\infty} \frac{2}{M} \sum_{\ell=1}^{M-1} \sum_{j_1,j_2=p}^{q} \varphi_{j_1}\varphi_{j_2} \cos\left(j_1\widetilde{\lambda}_\ell\right) \cos\left(-j_2\widetilde{\lambda}_\ell\right) \\
&\quad + \kappa_4 \left(\sum_{j=p}^{q} \varphi_j \lim_{n\to\infty} \frac{1}{M} \sum_{\ell=1}^{M-1} \cos\left(j\widetilde{\lambda}_\ell\right) \right)^2 \\
&= 2 \sum_{j_1,j_2=p}^{q} \varphi_{j_1}\varphi_{j_2} \int_0^1 \cos\left(\pi j_1 \lambda\right) \cos\left(-\pi j_2 \lambda\right) d\lambda + \kappa_4 \delta_j \varphi_0^2 \\
&= \sum_{j=p}^{q} \varphi_j^2 \left(1 + \delta_j + \kappa_4 \delta_j\right),
\end{aligned}$$

by elementary algebra and $g_{x,\ell}^{-1} g_x(\lambda_s) - 1 \to_{n\to\infty} 0$ by $C.1$ and $(2\ell-1)\, m < s \le (2\ell+1)\, m$, for the second equality that $\cos(z)\cos(w)$ is a differentiable function, see

Brillinger (1981, p.15) and that $M^{-1}\sum_{\ell=1}^{M-1}\cos\left(j\widetilde{\lambda}_\ell\right) \to \int_0^1 \cos(\pi j u)\, du = 0$ except for $j=0$ in which case is 1, whereas for the third equality we have used that the integral is 1 if $j_1 = j_2 = 0$, 2^{-1} if $j_1 = j_2 \neq 0$ and 0 otherwise. □

Proof of Corollary 1 By Delta methods and Theorem 1

$$n^{1/2}\left(\widehat{\sigma}_\varepsilon^2 - \sigma_{\varepsilon,n}^2\right) = n^{1/2}\left(\widehat{c}_0 - c_{0,n}\right)(2\pi)\, e^{c_{0,n}} + o_p(1).$$

But from Theorem 1 with $j=0$, $n^{1/2}\left(\widehat{c}_0 - c_{0,n}\right) \xrightarrow{d} N(0, 2+\kappa_4)$, whereas

$$(2\pi)\, e^{c_{0,n}} = (2\pi)\, e^{c_0} e^{c_{0,n}-c_0} = \sigma_\varepsilon^2 e^{c_{0,n}-c_0}$$

by (8). So, to complete the proof, it suffices to show that $e^{c_{0,n}-c_0} \to 1$. But by C.1

$$\begin{aligned} c_{0,n} - c_0 &= \frac{1}{M}\sum_{\ell=1}^{M}\log g_\ell - \int_0^1 \log g(\pi u)\, du \\ +\frac{1}{M}\sum_{\ell=1}^{M}\log\widetilde{\lambda}_\ell - \int_0^1 \log(\pi u)\, du &\to 0. \end{aligned}$$

by Brillinger (1980, p.15) and an obvious extension of Robinson's (1995) Lemma 2 respectively. □

Proof of Theorem 2 By Wold device, it suffices to examine the behaviour of

$$m^{1/2}\sum_{j=q}^{p}\varphi_j\left(\widehat{A}_j - A_{j,n}\right),$$

for any set of finite constants φ_j. Let $\widehat{d}_j = \log\left(\widehat{A}_j\right)$ and $\widetilde{d}_{j,n} = \log\left(\widetilde{A}_{j,n}\right)$ where $\widetilde{A}_{j,n} = \exp\left\{-\sum_{u=1}^{M-1}\widetilde{c}_{u,n}e^{-iu\widetilde{\lambda}_j}\right\}$.

We begin examining

$$\widehat{d}_j - d_{j,n} = -\sum_{u=1}^{M-1}\left(\widehat{c}_u - \widetilde{c}_{u,n}\right)e^{-iu\widetilde{\lambda}_j} - \sum_{u=1}^{M-1}\left(\widetilde{c}_{u,n} - c_{u,n}\right)e^{-iu\widetilde{\lambda}_j}, \tag{46}$$

where $d_{j,n} = \sum_{u=1}^{M-1} c_{u,n} e^{-iu\widetilde{\lambda}_j}$. By Taylor expansion of $\log\left(\widetilde{f}_{x,\ell}/f_{x,\ell}\right)$ the second term on the right of (46) is

$$\sum_{u=1}^{M-1} \frac{1}{M} \sum_{\ell=1}^{M} \left\{ \left(\frac{\widetilde{f}_{x,\ell} - f_{x,\ell}}{f_{x,\ell}} \right) + \left(\frac{\widetilde{f}_{x,\ell} - f_{x,\ell}}{f_{x,\ell}} \right)^2 (1 + o(1)) \right\} \cos\left(u\widetilde{\lambda}_\ell\right) e^{-iu\widetilde{\lambda}_j}$$

$$= \sum_{u=1}^{M-1} \frac{1}{M} \sum_{\ell=1}^{M} \left(\frac{\widetilde{f}_{x,\ell} - f_{x,\ell}}{f_{x,\ell}} \right) \cos\left(u\widetilde{\lambda}_\ell\right) e^{-iu\widetilde{\lambda}_j} + O\left(\frac{1}{M^3}\right)$$

$$= \frac{1}{2M} \sum_{\ell=1}^{M} \left(\frac{\widetilde{f}_{x,\ell} - f_{x,\ell}}{f_{x,\ell}} \right) \sum_{u=1}^{M-1} \left(e^{iu\widetilde{\lambda}_{\ell-j}} + e^{iu\widetilde{\lambda}_{\ell+j}} \right) + O\left(\frac{1}{M^3}\right) = O\left(\frac{\log M}{M^2}\right)$$

where for the first equality we have used Lemma 1 and in the second equality the proofs of Lemmas 1 and 4. So, by $C.4$ we conclude that it is $o\left(m^{-1/2}\right)$.

Next, the first term on the right of (46). From the definition of $\widehat{c}_u - \widetilde{c}_{u,n}$ in (41) and its properties in (42), this term is

$$-\sum_{u=1}^{M-1} \left(\frac{1}{M} \sum_{\ell=1}^{M-1} \left(\frac{\check{f}_{x,\ell} - \widetilde{f}_{x,\ell}}{\widetilde{f}_{x,\ell}} \right) \cos\left(u\widetilde{\lambda}_\ell\right) + \frac{1}{M} \sum_{\ell=1}^{M-1} \left(\frac{\widehat{f}_{x,\ell} - \check{f}_{x,\ell}}{\widetilde{f}_{x,\ell}} \right) \cos\left(u\widetilde{\lambda}_\ell\right) \right) e^{-iu\widetilde{\lambda}_j}$$
$$+ o_p\left(m^{-1/2}\right)$$
$$= -\sum_{u=1}^{M-1} \frac{1}{[n/2]} \sum_{\ell=1}^{[n/2]} \rho_\ell h_{u,n}(\ell) e^{-iu\widetilde{\lambda}_j}$$
$$-\frac{1}{2}(\widehat{\alpha} - \alpha) \sum_{u=1}^{M-1} \frac{1}{M} \sum_{\ell=1}^{M-1} g_{x,\ell} \left(\int_{-1}^{1} \log\left(1 + \frac{v}{2\ell}\right) dv \right) e^{-iu\widetilde{\lambda}_j} + o_p\left(m^{-1/2}\right),$$

by $C.4$ and Lemma 3, where $\rho_\ell = (2\pi)\sigma_\varepsilon^{-2} I_\varepsilon(\lambda_\ell) - 1$ and $h_{u,n}(\ell)$ was defined in Theorem 1. Thus,

$$m^{1/2}\left(\widehat{d}_j - \widetilde{d}_{j,n}\right) = -m^{1/2} \sum_{u=1}^{M-1} \left(\frac{1}{[n/2]} \sum_{\ell=1}^{[n/2]} \rho_\ell h_{u,n}(\ell) \right) e^{-iu\lambda_{2jm}} - \left(K \frac{1}{M} \sum_{u=1}^{M-1} e^{-iu\lambda_{2jm}} \right) m^{1/2}(\widehat{\alpha} - \alpha) + o_p(1). \quad (47)$$

since $\lim_{M\to\infty} \sum_{\ell=1}^{M-1} g_{x,\ell} \left(\int_{-1}^{1} \log\left(1 + \frac{v}{2\ell}\right) dv \right) = K$. But because $m = [n/(4M)]$, the first term on the right of (47) is

$$\frac{1}{2^{1/2}[n/2]^{1/2}} \sum_{\ell=1}^{[n/2]} \rho_\ell M^{-1/2} \sum_{u=1}^{M-1} h_{u,n}(\ell) e^{-iu\lambda_{2jm}} = \frac{1}{2^{1/2}[n/2]^{1/2}} \sum_{\ell=1}^{[n/2]} \rho_\ell \psi_{\ell,M(n)}(j)$$

where $\psi_{\ell,M(n)}(j) = M^{-1/2}\sum_{u=1}^{M-1} h_{u,n}(\ell)\, e^{-iu\lambda_{2jm}}$. Thus, $m^{1/2}\sum_{j=q}^{p}\varphi_j\left(\widehat{d}_j - \widetilde{d}_{j,n}\right)$ is

$$-\frac{1}{2^{1/2}[n/2]^{1/2}}\sum_{\ell=1}^{[n/2]}\rho_\ell\sum_{j=q}^{p}\varphi_j\psi_{\ell,M(n)}(j) - \frac{K}{2}\sum_{j=q}^{p}\varphi_j\left(\phi_j+\delta_j\right)m^{1/2}\left(\widehat{\alpha}-\alpha\right) + o_p(1). \tag{48}$$

Proceeding as in the proof of Theorem 1, the first term on the right of (48) converges in distribution to a complex normal random variable with variance

$$V = \lim_{n\to\infty}\sum_{j_1,j_2=q}^{p}\varphi_{j_1}\varphi_{j_2}\frac{1}{[n/2]}\sum_{\ell=1}^{[n/2]}\psi_{\ell,M(n)}(j_1)\,\overline{\psi_{\ell,M(n)}(j_2)}.$$

But, by definition of $h_{u,n}(\ell)$, the right side of the above equation is

$$\begin{aligned}
&\lim_{n\to\infty}\sum_{j_1,j_2=q}^{p}\varphi_{j_1}\varphi_{j_2}\frac{1}{M^2}\sum_{\ell=1}^{M-1}\sum_{u_1=1}^{M-1}\sum_{u_2=1}^{M-1}\left\{\cos\left(u_1\widetilde{\lambda}_\ell\right)\cos\left(-u_2\widetilde{\lambda}_\ell\right)e^{-iu_1\widetilde{\lambda}_{j_1}+iu_2\widetilde{\lambda}_{j_2}}\right\}\\
=\;& 2^{-1}\sum_{j_1,j_2=q}^{p}\varphi_{j_1}\varphi_{j_2}\left(\delta_{j_1-j_2}+2^{-1}\phi_{j_1}\phi_{j_2}-i\phi_{j_1-j_2}\right),
\end{aligned} \tag{49}$$

by Lemma 4. Next, by Robinson (1995) the second term on the right of (48) $\xrightarrow{d} N\left(0, 4^{-1}K^2\left|\sum_{j=q}^{p}\varphi_j\phi_j\right|^2\right)$, whereas the first and second terms of (48) are asymptotically independent since by Robinson (1995)

$$\left(\frac{1}{[n/2]^{1/2}}\sum_{\ell=1}^{[n/2]}\rho_\ell\right)m^{1/2}\left(\widehat{\alpha}-\alpha\right) = \frac{1}{[n/2]^{1/2}}\sum_{\ell=1}^{[n/2]}\rho_\ell\frac{1}{m^{1/2}}\sum_{\ell=1}^{m}v_\ell\rho_\ell + o_p(1),$$

where $v_\ell = \log\ell + m^{-1}\sum_{p=1}^{m}\log p$. But because by Brillinger's (1981) Theorem 5.2.4. $Cov\left(I_\varepsilon(\lambda_{\ell_1}), I_\varepsilon(\lambda_{\ell_2})\right) = O\left(n^{-1}\right)$, it implies that the expectation of the first term on the right is $o(1)$. So, from the behaviour of (48) and the second term on the right of (46),

$$m^{1/2}\sum_{j=q}^{p}\varphi_j\left(\widehat{d}_j - d_{j,n}\right)\xrightarrow{d} N(0,V),$$

where $V = \sum_{j_1,j_2=q}^{p}\varphi_{j_1}\varphi_{j_2}2^{-1}\left(\delta_{j_1-j_2}+2^{-1}\phi_{j_1}\phi_{j_2}-i\phi_{j_1-j_2}\right) + 4^{-1}K^2\left|\sum_{j=q}^{p}\varphi_j\phi_j\right|^2$.

But $\widehat{A}_j - A_{j,n} = \left(\exp\left(\widehat{d}_j - d_{j,n}\right)-1\right)A_{j,n}$ and $|A_{j,n}-A_j|\to 0$ from the proof of Lemma 7 which by a simple application of delta methods it implies that

$$m^{1/2}\sum_{j=q}^{p}\varphi_j\left(\widehat{A}_j - A_{j,n}\right)\xrightarrow{d} N\left(0,\sum_{j_1,j_2=q}^{p}\varphi_{j_1}\Omega_{A,j_1,j_2}\varphi_{j_2}\right),$$

and the proof is completed. $\square$

Proof of Theorem 3 Write $\widetilde{a}_{v,n} = (2M)^{-1} \sum_{j=-M+1}^{M} \widetilde{A}_{j,n} e^{iv\widetilde{\lambda}_j}$. The proof is completed if

$$(a)\;\; n^{1/2} \sum_{v=p}^{q} \varphi_v \left(\widehat{a}_v - \widetilde{a}_{v,n}\right) \xrightarrow{d} N\left(0, \sum_{v_1,v_2=p}^{q} \varphi_{v_1}\varphi_{v_2}\Omega_{a,v_1v_2}\right) \tag{50}$$

$$(b)\;\; n^{1/2} \sum_{v=p}^{q} \varphi_v \left(\widetilde{a}_{v,n} - a_{v,n}\right) \to 0.$$

We begin with (a). From the definition of $\widehat{a}_v - \widetilde{a}_{v,n}$ and Taylor expansion of $\widehat{A}_j - \widetilde{A}_{j,n}$, a typical element on the left of (50) is

$$\frac{n^{1/2}}{2M} \sum_{j=-M+1}^{M} \left(\widehat{d}_j - \widetilde{d}_{j,n}\right) \widetilde{A}_{n,j} e^{iv\widetilde{\lambda}_j} + \frac{n^{1/2}}{4M} \sum_{j=-M+1}^{M} \left|\widehat{d}_j - \widetilde{d}_{j,n}\right|^2 \left|\widetilde{A}_{n,j}\right| \left(1 + o_p(1)\right). \tag{51}$$

First, we show that the second term of (51) is $o_p(1)$. From (47) and trivial inequalities

$$\begin{aligned} m\left|\widehat{d}_j - \widetilde{d}_{j,n}\right|^2 &\leq G\left(\left|\frac{1}{[n/2]^{1/2}} \sum_{\ell=1}^{[n/2]} \rho_\ell \psi_{\ell,M(n)}(j)\right|^2 + \left|\phi_j m^{1/2}\left(\widehat{\alpha} - \alpha\right)\right|^2 + o_p(1)\right) \\ &= G\left|\frac{1}{[n/2]^{1/2}} \sum_{\ell=1}^{[n/2]} \rho_\ell \psi_{\ell,M(n)}(j)\right|^2 + O_p(1), \end{aligned}$$

where henceforth G is a generic finite positive constant. The contribution from the second term on the right of the last displayed equality into the second term of (51) is $O_p\left(m^{-1}n^{1/2}\right) = o_p(1)$ by $C.4$. Finally, the contribution from the first term on the right of the last displayed equality into the second term of (51) is also $o_p(1)$ by Theorem A in Serfling (1980, p.14) since by Theorem 2 and the continuous mapping theorem it converges to a χ_1^2 and by $C.2$ x_t^4 is uniformly integrable. So, from the definition of $\widehat{d}_j - \widetilde{d}_{j,n}$, (51) is

$$\begin{aligned} & -\frac{n^{1/2}}{2M} \sum_{j=-M+1}^{M} \left(\sum_{u=1}^{M-1} \left(\widehat{c}_u - \widetilde{c}_{u,n}\right) e^{-iu\widetilde{\lambda}_j}\right) \widetilde{A}_{n,j} e^{iv\widetilde{\lambda}_j} + O_p\left(\frac{n^{1/2}}{m}\right) \\ = & -\frac{n^{1/2}}{2M} \sum_{j=-M+1}^{M} \left(\sum_{u=1}^{M-1} \left(\frac{1}{M} \sum_{\ell=1}^{M-1} \left(\frac{\widehat{f}_{x,\ell} - \widetilde{f}_{x,\ell}}{\widetilde{f}_{x,\ell}}\right) \cos\left(u\widetilde{\lambda}_\ell\right)\right) e^{-iu\widetilde{\lambda}_j}\right) \widetilde{A}_{n,j} e^{iv\widetilde{\lambda}_j} + o_p(1) \\ = & -\frac{n^{1/2}}{M} \sum_{\ell=1}^{M-1} \left(\frac{\widehat{f}_{x,\ell} - \widetilde{f}_{x,\ell}}{\widetilde{f}_{x,\ell}}\right) \sum_{u=1}^{M-1} \cos\left(u\widetilde{\lambda}_\ell\right) \left(\frac{1}{2M} \sum_{j=-M+1}^{M} \widetilde{A}_{n,j} e^{i(v-u)\widetilde{\lambda}_j}\right) + o_p(1) \end{aligned}$$

$$= -\frac{n^{1/2}}{M}\sum_{\ell=1}^{M-1}\left(\frac{\widehat{f}_{x,\ell}-\widetilde{f}_{x,\ell}}{\widetilde{f}_{x,\ell}}\right)\sum_{u=1}^{M-1}\cos\left(u\widetilde{\lambda}_\ell\right)\left(\frac{1}{2M}\sum_{j=-M+1}^{M}A_j e^{i(v-u)\widetilde{\lambda}_j}\right)+o_p(1), \quad (52)$$

where for the first equality we have employed Taylor expansion and that by Hidalgo and Yajima's (1999) Proposition 3 $\sup_\ell\left|\widetilde{f}_{x,\ell}^{-1}\left(\widehat{f}_{x,\ell}-\widetilde{f}_{x,\ell}\right)\right| = o_p(1)$ and for the last equality that $M^{-1}\sum_{u=1}^{M-1}\left|\cos\left(u\widetilde{\lambda}_\ell\right)\right| = O\left(\ell^{-1}\right)$, $\left|\sum_{j=-M+1}^{M}\left(\widetilde{A}_{n,j}-A_j\right)e^{i(v-u)\widetilde{\lambda}_j}\right| = O(\log M)$ proceeding as with the proof of Lemma 7 and by Lemmas 2 and 3.

So,

$$n^{1/2}\sum_{v=p}^{q}\varphi_v\left(\widehat{a}_v-\widetilde{a}_{v,n}\right) = n^{1/2}\sum_{v=p}^{q}\varphi_v\left(\widehat{a}_v^1-\widetilde{a}_{v,n}\right)+\sum_{v=p}^{q}\psi_v$$

where $\psi_v = o_p(1)$ and

$$\begin{aligned}\widehat{a}_v^1-\widetilde{a}_{v,n} &= -\frac{1}{M}\sum_{\ell=1}^{M-1}\left(\frac{\widehat{f}_{x,\ell}-\widetilde{f}_{x,\ell}}{\widetilde{f}_{x,\ell}}\right)\sum_{u=1}^{M-1}\cos\left(u\widetilde{\lambda}_\ell\right)\left(\frac{1}{2M}\sum_{j=-M+1}^{M}A_j e^{i(v-u)\widetilde{\lambda}_j}\right) \quad (53)\\ &= -\frac{1}{M}\sum_{\ell=1}^{M-1}\left(\frac{\check{f}_{x,\ell}-\widetilde{f}_{x,\ell}}{\widetilde{f}_{x,\ell}}\right)\sum_{u=1}^{M-1}\cos\left(u\widetilde{\lambda}_\ell\right)\left(\frac{1}{2M}\sum_{j=-M+1}^{M}A_j e^{i(v-u)\widetilde{\lambda}_j}\right)+n^{-1/2}h_v,\end{aligned}$$

proceeding as in the proofs of Theorems 1 and 2 where $h_v = o_p(1)$. Now use these two Theorems to conclude that

$$n^{1/2}\sum_{v=p}^{q}\varphi_v\left(\widehat{a}_v-\widetilde{a}_{v,n}\right)\xrightarrow{d}N(0,V)$$

where, denoting $(2M)^{-1}\sum_{j=-M+1}^{M}A_j e^{i(v-u)\widetilde{\lambda}_j} = a_{v-u}^*$,

$$\begin{aligned}V &= \lim_{M\to\infty}\frac{1}{M}\sum_{\ell=1}^{M-1}\left(\sum_{v=p}^{q}\varphi_v\sum_{u=1}^{M-1}\cos\left(u\widetilde{\lambda}_\ell\right)a_{v-u}^*\right)^2\\ &= \sum_{v_1,v_2=p}^{q}\varphi_{v_1}\varphi_{v_2}\lim_{M\to\infty}\frac{1}{M}\sum_{u_1,u_2=1}^{M-1}\sum_{\ell=1}^{M-1}\cos\left(u_1\widetilde{\lambda}_\ell\right)a_{v_1-u_1}^*\cos\left(u_2\widetilde{\lambda}_\ell\right)a_{v_2-u_2}^*.\end{aligned}$$

A typical component of the last displayed equation is

$$\begin{aligned}&\lim_{M\to\infty}\frac{1}{2M}\sum_{u_1,u_2=1}^{M-1}\sum_{\ell=1}^{M-1}a_{v_1-u_1}^*a_{v_2-u_2}^*\left(\cos\left((u_1+u_2)\widetilde{\lambda}_\ell\right)+\cos\left((u_1-u_2)\widetilde{\lambda}_\ell\right)\right)\\ =\ &\lim_{M\to\infty}\left(\frac{1}{2}\sum_{u=1}^{M-1}a_{v_1-u}^*a_{v_2-u}^*-\frac{1}{M}\sum_{\substack{u_1=1,\neq u_2\\ u_1\pm u_2=odd}}^{M-1}a_{v_1-u_1}^*a_{v_2-u_2}^*\right)\end{aligned}$$

using (65) in Lemma 4 for the second term on the right of the last equation. But using $a_u = O\left(u^{-1-\alpha}\right)$ and that since $A(\lambda)$ has an integrable derivative by Brillinger (1981, p.15) $a^*_{v-u} = a_{u-v} + O\left(M^{-1}\right)$, we conclude that

$$V = \frac{1}{2}\sum_{v_1,v_2=p}^{q} \varphi_{v_1}\varphi_{v_2}\sum_{u=1}^{\infty} a_{v_1-u}a_{v_2-u} = \sum_{v_1,v_2=p}^{q} \varphi_{v_1}\varphi_{v_2}\Omega_{a,v_1v_2},$$

which completes the proof of part (a). Now part (b) follows by identical arguments to those in (52) and Lemma 1. □

Proof of Theorem 4 (a) From the definition of $\widehat{x}^*_n$ and $\widetilde{x}_n$, their difference is

$$\varepsilon_t - \sum_{u=M}^{\infty} a_u\widetilde{x}_{n-u} + \sum_{u=1}^{M-1}\left(\widehat{a}_u - a_{u,n}\right)\widetilde{x}_{n-u} + \sum_{u=1}^{M-1}\left(a_{u,n} - a_u\right)\widetilde{x}_{n-u}. \qquad (54)$$

The second moment of the second term of (54) is

$$\sigma^2_{\widetilde{x}}\sum_{u=M}^{\infty} a_u^2 + 2\sum_{M=u_1<u_2}^{\infty} a_{u_1}a_{u_2}\gamma_{\widetilde{x}}\left(u_1-u_2\right) = O\left(M^{-1}\right).$$

because $a_u = O(u^{-1-\alpha/2})$ and $\gamma_{\widetilde{x}}(u) \sim Cu^{-1+\alpha}$, $|C| < \infty$ by C.1 and C.2. Next, we examine the third term of (54). Denoting the first term on the right of (53) $n^{1/2}q_u$, the third term of (54) is

$$\sum_{u=1}^{M-1} q_u\widetilde{x}_{n-u} + \sum_{u=1}^{M-1}\left(\widehat{a}_u - a_{u,n} - q_u\right)\widetilde{x}_{n-u}.$$

But $\sup_{u<M}\left|\widehat{a}_u - a_{u,n} - q_u\right| \leq G\sum_{u=1}^{M-1}\left|\widehat{a}_u - a_{u,n} - q_u\right| = O_p\left(Mm^{-1}\right)$ since $\left|\widehat{a}_u - a_{u,n} - q_u\right| = O_p\left(m^{-1}\right)$, so that the second term of the last displayed expression is $O_p\left(M^2m^{-1}\right) = o_p(1)$ by C.4. On the other hand, the second moment of first term is

$$n^{-1}\left(\sigma^2_{\widetilde{x}}\sum_{u=1}^{M-1} E\left(n^{1/2}q_u\right)^2 + 2\sum_{1=u_1<u_2}^{M-1} E\left(nq_{u_1}q_{u_2}\right)\gamma_{\widetilde{x}}\left(u_1-u_2\right)\right) = o\left(M^{\alpha-1}\right),$$

after elementary algebra by the Cauchy-Schwarz inequality, Theorem 3 and Serfling (1980, p.14) by the uniform integrability of $n^{1/2}q_u$. Finally, the second moment of the fourth term of (54) is $O\left(M^{\alpha-1}\log M\right)$ by similar arguments since by Lemma 7

$(a_{u,n} - a_u) = O(M^{-1}\log M)$. From here, the conclusion is immediate since $\alpha < 1$ and $E(\varepsilon_t^2) = \sigma_\varepsilon^2$.

(b) By definition

$$\widetilde{x}_{n+h} = \sum_{j=0}^{h} b_j \varepsilon_{t+h-j} + \sum_{u=h+1}^{\infty}\left(\sum_{\ell=0}^{h} b_\ell a_{u-\ell}\right)\widetilde{x}_{t+h-u},$$

where b_j, $j = 0, 1, ...$, satisfies $\left(\sum_{j=0}^{\infty} b_j L^j\right)\left(\sum_{j=0}^{\infty} a_j L^j\right) = 1$. Thus,

$$\begin{aligned}\widetilde{x}_{n+h} - \widehat{x}^*_{n+h} &= \sum_{j=0}^{h} b_j \varepsilon_{t+h-j} + \sum_{u=h+1}^{M}\left[\left(\sum_{\ell=0}^{h} b_\ell a_{u-\ell}\right) - \left(\sum_{\ell=0}^{h} \widehat{b}_\ell \widehat{a}_{u-\ell}\right)\right]\widetilde{x}_{t+h-u} \\ &\quad + \sum_{u=M+1}^{\infty}\left(\sum_{\ell=0}^{h} b_\ell a_{u-\ell}\right)\widetilde{x}_{t+h-u}.\end{aligned}$$

Now, by similar arguments to those used for the last two terms of (54), the last two terms on the right of the last equation are $o_p(1)$. So,

$$\widetilde{x}_{n+h} - \widehat{x}^*_{n+h} = \sum_{j=0}^{h} b_j \varepsilon_{t+h-j} + o_p(1).$$

From here the conclusion follows by $C.2$. □

Proof of Theorem 5 From the definition of $\widehat{\psi}_j$ in (24), it follows that

$$\widehat{\psi}_j = \frac{\widehat{\sigma}^2_{\varepsilon^x}}{\widehat{\sigma}^2_\varepsilon}\sum_{p=0}^{N-|j|}\widehat{\beta}_p\widehat{\beta}_{p+|j|}, \quad \text{with } \widehat{\beta}_0 = 1.$$

Since by Corollary 1, Robinson (1995) and Slutzky Theorem $\widehat{\sigma}_\varepsilon^{-2}\widehat{\sigma}^2_{\varepsilon^x} - \sigma_\varepsilon^{-2}\sigma^2_{\varepsilon^x} = O_p\left(m^{-1/2}\right)$, it suffices to show that

$$\sum_{j=\max(t-n,1-N)}^{\min(t-1,N-1)}\left(\sum_{p=0}^{N-|j|}\widehat{\beta}_p\widehat{\beta}_{p+|j|}\right)y_{t-j} - \sum_{j=-\infty}^{\infty}\psi_j y_{t-j} = o_p(1).$$

The left side of the last displayed equation is

$$\sum_{j=\max(t-n,1-N)}^{\min(t-1,N-1)}\sum_{p=0}^{N-|j|}\left(\widehat{\beta}_p\widehat{\beta}_{p+|j|} - \widetilde{\beta}_{n,p}\widetilde{\beta}_{n,p+|j|}\right)y_{t-j} \tag{55}$$

$$+\sum_{j=\max(t-n,1-N)}^{\min(t-1,N-1)}\left(\sum_{p=0}^{N-|j|}\widetilde{\beta}_{n,p}\widetilde{\beta}_{n,p+|j|}-\psi_j\right)y_{t-j} \tag{56}$$

$$+\left\{\sum_{j<\max(t-n,1-N)}+\sum_{j>\min(t-1,N-1)}\right\}\psi_j y_{t-j} \tag{57}$$

where $\widetilde{\beta}_{n,p}$ is defined as $\widehat{\beta}_p$ but with $\widehat{g}(\lambda)$ replaced by $g(\lambda)$. The second moment of the second term of (57) is

$$\sigma_y^2\sum_{j=\min(t-1,N-1)}^{\infty}\psi_j^2+2\sum_{\min(t-1,N-1)=j<k}^{\infty}\psi_j\psi_k\gamma_{k-j}=O\left(N^{-1-\alpha}+N^{-1}\right)=o(1)$$

by $C.5$, because $\gamma_{k-j}=O\left(|k-j|^{\alpha-1}\right)$, $\psi_j=O\left(j^{-1-\alpha/2}\right)$, $N\to\infty$ and $t=[n\delta]$ for any arbitrary $\delta\in(0,1)$. The second moment of the first term of (57) follows similarly.

Next, since $\psi_j=\sum_{p=0}^{\infty}\beta_p\beta_{p+|j|}$, it implies that $\phi_j=\sum_{p=0}^{N-|j|}\widetilde{\beta}_{n,p}\widetilde{\beta}_{n,p+|j|}-\psi_j$ is

$$\sum_{p=0}^{N-|j|}\left(\widetilde{\beta}_{n,p}\widetilde{\beta}_{n,p+|j|}-\beta_p\beta_{p+|j|}\right)-\sum_{p=N-|j|+1}^{\infty}\beta_p\beta_{p+|j|}=O\left(N^{-1}\log N\right)$$

by Lemma 7 once we identify $a_{u,n}$ and a_u as $\widetilde{\beta}_{n,p}$ and β_p respectively and since $\beta_p=o\left(p^{-1}\right)$ as can easily be seen from the definition of $g(\lambda)$. So, proceeding as above, the second moment of (56) is

$$\sigma_y^2\sum_{j=\max(t-n,1-N)}^{\min(t-1,N-1)}\phi_j^2+2\sum_{\max(t-n,1-N)=j<k}^{\min(t-1,N-1)}\phi_j\phi_k\gamma_{k-j}=O\left(N^{-2}N^{1+\alpha}\log M\right)=o(1)$$

because $\gamma_j=O\left(j^{\alpha-1}\right)$, $0<\alpha<1$ and $C.5$.

Finally (55). Define $\breve{\beta}_p$ as $\widehat{\beta}_p$ but with $\widehat{\alpha}$ replaced by α, that is in the formulae (13) – (15) we employ

$$\breve{g}(\lambda)=\frac{1}{2m+1}\sum_{\ell=-m}^{m}|\lambda_\ell+\lambda|^{\alpha}I_y(\lambda_\ell+\lambda)$$

instead of $\widehat{g}(\lambda)$. Since by Robinson (1995) $\widehat{\alpha}-\alpha=O_p\left(m^{-1/2}\right)$, then

$$\left|\sum_{j=\max(t-n,1-N)}^{\min(t-1,N-1)}\sum_{p=0}^{N-|j|}\left(\widehat{\beta}_p\widehat{\beta}_{p+|j|}-\breve{\beta}_p\breve{\beta}_{p+|j|}\right)y_{t-j}\right|$$

$$\leq\sum_{p=0}^{N}\left|\widehat{\beta}_p\widehat{\beta}_{p+|j|}-\breve{\beta}_p\breve{\beta}_{p+|j|}\right|\sum_{j=\max(t-n,1-N)}^{\min(t-1,N-1)}|y_{t-j}|=O_p\left(Nm^{-1/2}\right)=o_p(1)$$

by $C.5$. So, to complete the proof we need to show that

$$\sum_{j=\max(t-n,1-N)}^{\min(t-1,N-1)} \sum_{p=0}^{N-|j|} \left(\check{\beta}_p \check{\beta}_{p+|j|} - \widetilde{\beta}_{n,p} \widetilde{\beta}_{n,p+|j|} \right) y_{t-j} = o_p(1). \tag{58}$$

But in Theorem 3, denoting $B_j = \sum_{q=0}^{\infty} \beta_q e^{iq\widetilde{\lambda}_j}$, we have shown, c.f. (53), that

$$\check{\beta}_p - \widetilde{\beta}_{n,p} = -\frac{1}{M} \sum_{\ell=1}^{M-1} \left(\frac{\check{g}_\ell - \widetilde{g}_\ell}{\widetilde{g}_\ell} \right) \sum_{u=1}^{M-1} \cos\left(u \widetilde{\lambda}_\ell \right) \left(\frac{1}{2M} \sum_{j=-M+1}^{M} B_j e^{i(p-u)\widetilde{\lambda}_j} \right) + n^{-1/2} h_p$$

where $\widetilde{g}_\ell = (2m+1)^{-1} \sum_{q=-m}^{m} |\lambda_\ell + \lambda_q|^\alpha g(\lambda_\ell + \lambda_q)$. Denoting $\overline{\beta}_p$ the first term on the right of the last displayed equation, the left side of (58) is, except negligible cross-terms,

$$\sum_{j=\max(t-n,1-N)}^{\min(t-1,N-1)} \check{\phi}_{j,n} y_{t-j} + \sum_{j=\max(t-n,1-N)}^{\min(t-1,N-1)} \overline{h}_j y_{t-j}, \tag{59}$$

where $\overline{h}_j$ and $\check{\phi}_{j,n}$ involve the contribution of $N - |j|$ terms of the type $n^{-1} h_\ell h_k$ and $\overline{\beta}_\ell \overline{\beta}_k$ respectively, with $\ell - k = j$, so that $E\check{\phi}_{j,n}^2 = O(N^2 n^{-2})$ since by Theorem 3 $E\left|\overline{\beta}_p\right|^2 = O(n^{-1})$. Now, the first absolute moment of the first term on the right of (59) is

$$\begin{aligned}
\sum_{j=\max(t-n,1-N)}^{\min(t-1,N-1)} E\left|\check{\phi}_{j,n} y_{t-j}\right| &\leq \sum_{j=\max(t-n,1-N)}^{\min(t-1,N-1)} \left(E\check{\phi}_{j,n}^2\right)^{1/2} \left(E y_{t-j}^2\right)^{1/2} \\
&= O\left(N^2 n^{-1}\right) = o(1),
\end{aligned}$$

by $C.5$ and the Cauchy-Schwarz inequality since $E\left(y_{t-j}^2\right) \leq G$. Next, by the Cauchy-Schwarz inequality the contribution of the second term of (59) is bounded by

$$\left(\sum_{j=\max(t-n,1-N)}^{\min(t-1,N-1)} \overline{h}_j^2 \right)^{1/2} \left(\sum_{j=\max(t-n,1-N)}^{\min(t-1,N-1)} y_{t-j}^2 \right)^{1/2} = O_p\left(N^2 n^{-1}\right)$$

because $\overline{h}_j = O_p\left(n^{-1/2}\right)$. So, by $C.4$, $C.5$ and Markov's inequality (59) is $o_p(1)$, which completes the proof. □

7. TECHNICAL LEMMAS

Lemma 1 *Assuming C.1,* $\widetilde{c}_{u,n} - c_{u,n} = O\left(M^{-2}\right)$.

Proof. Since from the definition of $\widetilde{f}_{x,\ell}$ and Taylor expansion of $\log(z)$,

$$\widetilde{c}_{u,n} - c_{u,n} = \frac{1}{M}\sum_{\ell=1}^{M}\left(\frac{\widetilde{f}_{x,\ell} - f_{x,\ell}}{f_{x,\ell}}\right)\cos\left(u\widetilde{\lambda}_\ell\right) + \frac{1}{M}\sum_{\ell=1}^{M}\left(\frac{\widetilde{f}_{x,\ell} - f_{x,\ell}}{f_{x,\ell}}\right)^2\cos\left(u\widetilde{\lambda}_\ell\right)(1 + o(1)),$$

the proof is completed if we show that $f_{x,\ell}^{-1}\left(\widetilde{f}_{x,\ell} - f_{x,\ell}\right) = O\left(M^{-2}\right)$ uniformly in ℓ. But,

$$\begin{aligned} f_{x,\ell}^{-1}\left(\widetilde{f}_{x,\ell} - f_{x,\ell}\right) &= \frac{g_{x,\ell}^{-1}}{(2m+1)}\sum_j\left(g_x\left(\lambda_{j+2m\ell}\right) - g_x\left(\lambda_{2m\ell}\right)\right) \\ &= \frac{g_{x,\ell}^{-1}}{(2m+1)}\sum_j\left(\frac{2\pi j}{n}\right)^2\frac{\partial^2}{\partial\lambda^2}g_x\left(\lambda_{2m\ell} + \xi\lambda_j\right) \\ &= g_{x,\ell}^{-1}\frac{\partial^2}{\partial\lambda^2}\left(g_x\left(\lambda_{2m\ell}\right)\right)\frac{\pi^2}{4M^2}\frac{1}{(2m+1)}\sum_j\left(\frac{j}{m}\right)^2(1 + o(1)), \end{aligned}$$

where $\xi = \xi(j) \in (0,1)$ and because $\frac{\partial^2}{\partial\lambda^2}g_x\left(\lambda_{2m\ell} + \lambda_j\right) = \frac{\partial^2}{\partial\lambda^2}g_{x,\ell}(1 + o(1))$ for all j by $C.1$. The proof now follows since $(2m+1)^{-1}\sum_j\left(\frac{j}{m}\right)^2 \leq 1$. □

Lemma 2 *Assuming C.1-C.4, for all* $\ell = 1, 2, \ldots$

$$E\left(\widetilde{f}_{x,\ell}^{-1}\left(\check{f}_{x,\ell} - \widetilde{f}_{x,\ell}\right)\right)^2 = O\left(m^{-1/2}\right).$$

Proof. By definition of $\widetilde{f}_{x,j}$ and $\check{f}_{x,j}$, the left side of the last displayed equality is

$$\left((2m+1)^{-1}\sum_j g_x^{-1}\left(\lambda_{j+2m\ell}\right)\right)^{-2} E\left((2m+1)^{-1}\sum_j g_x\left(\lambda_{j+2m\ell}\right)\left(\frac{I\left(\lambda_{j+2m\ell}\right)}{f_x\left(\lambda_{j+2m\ell}\right)} - 1\right)\right)^2.$$

But, the first factor by $C.1$ is bounded and bounded away from zero whereas the second factor is $O\left(m^{-1}\right)$ by routine extension of Robinson (1995) since $g_x\left(\lambda_{j+2m\ell}\right)$ is a differentiable function by $C.1$. □

Lemma 3 *Assuming C.1-C.4,*

$$\begin{aligned} \widetilde{f}_{x,\ell}^{-1}\left(\widehat{f}_{x,\ell} - \check{f}_{x,\ell}\right) &= \frac{\widehat{\alpha} - \alpha}{2}g_{x,\ell}\left(\int_{-1}^{1}\log\left(1 + \frac{v}{2\ell}\right)dv\right) \\ &\quad + O_p\left(m^{-1}\log^2\left(\lambda_{2m\ell}\right) + m^{-1/2}M^{-1}\ell^{-1}\right). \end{aligned} \tag{60}$$

Proof. By definition of $\widetilde{f}_{x,\ell}$, $\widehat{f}_{x,\ell}$ and $\check{f}_{x,\ell}$ and elementary algebra, the left side of (60) is

$$\widetilde{g}_x^{-1}\left(\lambda_{2m\ell}\right)\left\{\left(\lambda_{2m\ell}^{-(\widehat{\alpha}-\alpha)}-1\right)\frac{1}{(2m+1)}\sum_j \lambda_{2m\ell+j}^{\alpha} I\left(\lambda_{j+2m\ell}\right)\left(\lambda_{2m\ell+j}^{(\widehat{\alpha}-\alpha)}-1\right)\right. \quad (61)$$
$$\left.+\frac{1}{(2m+1)}\sum_j \lambda_{2m\ell+j}^{\alpha} I\left(\lambda_{j+2m\ell}\right)\left\{\left(\lambda_{2m\ell}^{-(\widehat{\alpha}-\alpha)}-1\right)+\left(\lambda_{2m\ell+j}^{(\widehat{\alpha}-\alpha)}-1\right)\right\}\right\}.$$

where $\widetilde{g}_x\left(\lambda_{2m\ell}\right) = (2m+1)^{-1}\sum_j g_x\left(\lambda_{j+2m\ell}\right)$. By $C.1$ the first factor in (61) is bounded uniformly in ℓ. Next

$$\begin{aligned}\frac{1}{(2m+1)}\sum_j \lambda_{2m\ell+j}^{\alpha} I\left(\lambda_{j+2m\ell}\right) &= \frac{1}{(2m+1)}\sum_j g_x\left(\lambda_{2m\ell+j}\right)\left(\frac{I\left(\lambda_{j+2m\ell}\right)}{f_x\left(\lambda_{j+2m\ell}\right)}-1\right)\\ &\quad+\frac{1}{(2m+1)}\sum_j g_x\left(\lambda_{2m\ell+j}\right) \qquad (62)\\ &= g_{x,\ell}+O\left(M^{-2}\right)+O_p\left(m^{-1/2}\right),\end{aligned}$$

by Lemma 2 and Markov's inequality since by $C.1$ $g_x(\lambda)$ is twice continuously differentiable, $\sum_j j = 0$ and Robinson (1995).

So, the first term inside the curly brackets of (61) is $O_p\left(m^{-1}\log^2\lambda_{2m\ell}\right)$ since by Robinson (1995) $(\widehat{\alpha}-\alpha) = O_p\left(m^{-1/2}\right)$, $\left|\lambda_{2m\ell+j}^{-(\widehat{\alpha}-\alpha)}-1\right| \leq |\widehat{\alpha}-\alpha|\,|\log\lambda_{2m\ell+j}|$, for $j = 0, \pm 1, ...$, and (62). So, we are left to examine the contribution from the second term inside the curly brackets of (61), which by a Taylor expansion is

$$\begin{aligned}&\frac{(\widehat{\alpha}-\alpha)}{(2m+1)}\sum_j \lambda_{2m\ell+j}^{\alpha} I\left(\lambda_{j+2m\ell}\right)\left(\log\left(\lambda_{2m\ell+j}\right)-\log\left(\lambda_{2m\ell}\right)\right)+O_p\left(m^{-1}\log^2\lambda_{2m\ell}\right)\\ =\ &(\widehat{\alpha}-\alpha)\,g_{x,\ell}\frac{1}{(2m+1)}\sum_j \log\left(1+\frac{j}{2m\ell}\right)+O_p\left(m^{-1}\log^2\lambda_{2m\ell}+m^{-1/2}M^{-1}\ell^{-1}\right)\\ =\ &\frac{1}{2}(\widehat{\alpha}-\alpha)\,g_{x,\ell}\int_{-1}^{1}\log\left(1+\frac{v}{2\ell}\right)dv+O_p\left(m^{-1}\log^2\lambda_{2m\ell}+m^{-1/2}M^{-1}\ell^{-1}\right),\end{aligned}$$

proceeding as with (62) and using that $\log\left(1+j/\left(2m\ell\right)\right) = O\left(\ell^{-1}\right)$ and $g_x\left(\lambda_{2m\ell+j}\right) - g_x\left(\lambda_{2m\ell}\right) = O\left(M^{-1}\right)$, for the first equality and for the second one that by Brillinger (1981, p.15)

$$\left|(2m+1)^{-1}\sum_j \log\left(1+\frac{j}{2m\ell}\right)-\frac{1}{2}\int_{-1}^{1}\log\left(1+\frac{v}{2\ell}\right)dv\right| \leq Km^{-1}. \qquad \square$$

Lemma 4

$$\frac{1}{M^2}\sum_{p=1}^{M}\left(\sum_{u_1=1}^{M}\cos\left(u_1\widetilde{\lambda}_p\right)e^{-iu_1\widetilde{\lambda}_{j_1}}\right)\left(\sum_{u_2=1}^{M}\cos\left(-u_2\widetilde{\lambda}_p\right)e^{iu_2\widetilde{\lambda}_{j_2}}\right) \tag{63}$$
$$=2^{-1}\left(\delta_{j_1-j_2}+2^{-1}\phi_{j_1}\phi_{j_2}-i\phi_{j_1-j_2}\right)+O\left(M^{-1}\right).$$

Proof. Because $\cos x\cos y=(\cos(x-y)+\cos(x+y))/2$, the left side of (63) is

$$\frac{1}{2M}\sum_{u_1=1}^{M}\sum_{u_2=1}^{M}e^{-iu_1\widetilde{\lambda}_{j_1}}e^{iu_2\widetilde{\lambda}_{j_2}}\frac{1}{M}\sum_{p=1}^{M}\left(\cos\left(\left(u_1+u_2\right)\widetilde{\lambda}_p\right)+\cos\left(\left(u_1-u_2\right)\widetilde{\lambda}_p\right)\right). \tag{64}$$

But, since $\widetilde{\lambda}_p=(\pi p)/M$, the only terms in (64) different than zero are when $u_1=u_2$, or $u_1\pm u_2=2\ell+1$ for some integer ℓ. When $u_1=u_2$ (64) is

$$\frac{1}{2M}\sum_{u=1}^{M}e^{-iu\widetilde{\lambda}_{j_1-j_2}}=\frac{1}{2}\left(\delta_{j_1-j_2}-\frac{i}{\pi\left(j_1-j_2\right)}\left(1-\cos\left(\left(j_2-j_1\right)\pi\right)\right)\right)+O\left(\frac{1}{M}\right),$$

using standard Euler-MacLaurin approximation of sums by integrals. Next, using formulae in Brillinger (1981, p.13) for $u_1+u_2=2\ell_1+1$ or $u_1-u_2=2\ell_2+1$

$$\sum_{p=1}^{M}\left(\cos\left(\left(u_1+u_2\right)\widetilde{\lambda}_p\right)+\cos\left(\left(u_1-u_2\right)\widetilde{\lambda}_p\right)\right)$$
$$=\quad -1+\frac{\sin\left(M\widetilde{\lambda}_{u_1+u_2}+\frac{1}{2}\widetilde{\lambda}_{u_1+u_2}\right)}{2\sin\left(\frac{1}{2}\widetilde{\lambda}_{u_1+u_2}\right)}+\frac{\sin\left(M\widetilde{\lambda}_{u_1-u_2}+\frac{1}{2}\widetilde{\lambda}_{u_1-u_2}\right)}{2\sin\left(\frac{1}{2}\widetilde{\lambda}_{u_1-u_2}\right)}=-2 \tag{65}$$

since $\sin\left(\left(2\ell+1\right)\pi+\lambda\right)=-\sin\left(\lambda\right)$. So, for $u_1+u_2=2\ell_1+1$ or $u_1-u_2=2\ell_2+1$ (64) is

$$-\frac{1}{M^2}\sum_{u_1,u_2=1;u_1\pm u_2=odd}^{M}e^{-iu_1\widetilde{\lambda}_{j_1}}e^{iu_2\widetilde{\lambda}_{j_2}}. \tag{66}$$

Consider the situation that M is odd. In this case (66) is

$$-\frac{1}{M^2}\sum_{\ell_1=0}^{[M/2]}\sum_{\ell_2=1}^{[M/2]}e^{-i(2\ell_1+1)\widetilde{\lambda}_{j_1}}e^{i2\ell_2\widetilde{\lambda}_{j_2}}-\frac{1}{M^2}\sum_{\ell_1=1}^{[M/2]}\sum_{\ell_2=0}^{[M/2]}e^{-i2\ell_1\widetilde{\lambda}_{j_1}}e^{i(2\ell_2+1)\widetilde{\lambda}_{j_2}}$$
$$=\quad -\left(e^{-i\widetilde{\lambda}_{j_1}}+e^{i\widetilde{\lambda}_{j_2}}\right)\frac{1}{M^2}\left(\sum_{\ell_1=1}^{[M/2]}e^{-i2\ell_1\widetilde{\lambda}_{j_1}}\right)\left(\sum_{\ell_2=1}^{[M/2]}e^{i2\ell_2\widetilde{\lambda}_{j_2}}\right)$$
$$=\quad \frac{1}{2\pi j_1}\left(1-\cos\left(\pi j_1\right)\right)\frac{1}{2\pi j_2}\left(1-\cos\left(\pi j_2\right)\right)\left(1+O\left(\frac{1}{M}\right)\right).$$

Next, when M is even (66) is

$$\begin{aligned} & -\frac{1}{M^2} \sum_{u_1,u_2=1;u_1 \pm u_2=odd}^{M-1} e^{-iu_1\widetilde{\lambda}_{j_1}} e^{iu_2\widetilde{\lambda}_{j_2}} + O\left(\frac{1}{M}\right) \\ = & \frac{1}{2\pi j_1}\left(1-\cos\left(\pi j_1\right)\right) \frac{1}{2\pi j_2}\left(1-\cos\left(\pi j_2\right)\right)\left(1+O\left(\frac{1}{M}\right)\right) \end{aligned}$$

proceeding as when M is odd. From here the conclusion of the lemma is obvious. □

Lemma 5 *Denoting by* $\mathcal{I}(\cdot)$ *the indicator function,*

$$\sum_{s=M}^{\infty} c_s e^{-is\lambda_{2mj}} = O\left(\lambda_{2mj}^{-1} M^{-1} \mathcal{I}(\alpha>0) + M^{-1}\mathcal{I}(\alpha=0)\right). \tag{67}$$

Proof. From the definition of c_s in (10) and $C.1$, $c_s = c_{s1} + c_{s2}$ where

$$c_{s1} = -\frac{\alpha}{\pi}\int_0^{\pi} (\log\lambda)\cos(s\lambda)\,d\lambda \text{ and } c_{s2} = \frac{1}{\pi}\int_0^{\pi}\log(g(\lambda))\cos(s\lambda)\,d\lambda.$$

Since $g(\lambda)$ is twice continuously differentiable function by $C.1$ $c_{s2} = O(s^{-3})$, whereas by integration by parts c_{s1} is

$$-\frac{\alpha}{\pi s}\sin(s\lambda)\log\lambda|_0^{\pi} + \frac{\alpha}{\pi s}\int_0^{\pi}\frac{\sin(s\lambda)}{\lambda}d\lambda = \frac{\alpha}{2s} + O\left(s^{-2}\right)$$

since $\int_0^{\infty}\lambda^{-1}\sin(\lambda)\,d\lambda = \pi/2$ and as $s\to\infty$ by Courant and John (1974, sect. 8.4.c)

$$\left|\int_0^{\pi}\frac{\sin(s\lambda)}{\lambda}d\lambda - \int_0^{\infty}\frac{\sin(\lambda)}{\lambda}d\lambda\right| \le Ks^{-1}.$$

Thus, the left side of (67) is bounded in absolute value by

$$\begin{aligned} \sum_{s=M}^{\infty} c_s e^{-is\lambda_{2mj}} &= \sum_{s=M}^{\infty} c_{s2} e^{-is\lambda_{2mj}} + \sum_{s=M}^{\infty} c_{s1} e^{-is\lambda_{2mj}} \\ &= O\left(M^{-1}\right) + \frac{\alpha}{2}\sum_{s=M}^{\infty} s^{-1} e^{-is\lambda_{2mj}}, \end{aligned}$$

whose second term on the right by Abel summation by parts is

$$\alpha\left|\sum_{s=M}^{\infty}\left(s^{-1}-(s+1)^{-1}\right)\sum_{\ell=M}^{s} e^{-i\ell\lambda_{2mj}}\right| \le G\alpha\lambda_{2mj}^{-1}\sum_{s=M}^{\infty} s^{-2}.$$

From here the conclusion of the lemma is obvious. □

Lemma 6 *Let* $\mathcal{B}_1(\pi M\lambda) = \pi M\lambda - [\pi M\lambda] - 1/2$ *the first Bernstein polynomial. Then*

$$(a)\quad \int_0^1 \frac{\cos(x\pi u) - 1}{x\pi}\mathcal{B}_1(Mx\pi)\,dx = O\left(M^{-1}\right) \quad (68)$$

$$(b)\quad \int_0^1 \log(x\pi)\sin(x\pi u)\mathcal{B}_1(Mx\pi)\,dx = O\left(M^{-2}\log M + g_1(u) - g_2(u)\right), \quad (69)$$

where $u \le M$, $g_1(u) = -(4\pi)^{-1} u \sum_{v=1}^{\infty} v^{-1}\left(4M^2\pi^2v^2 - u^2\right)^{-1}$ *and*
$g_2(u) = \left(\log(\pi)/\left(2\pi^2\right)\right) u(-1)^{u+1}\sum_{v=1}^{\infty} v^{-1}\sin\left(2M\pi^2 v\right)\left(4M^2\pi^2v^2 - u^2\right)^{-1}$.

Proof. We begin with (a). By a change of variables, the left side of (68) is

$$\begin{aligned}\frac{1}{\pi}\int_0^{M\pi}\frac{\cos\left(\frac{tu}{M}\right)-1}{t}\mathcal{B}_1(t)\,dt &= \frac{1}{\pi}\sum_{\ell=1}^{M}\int_{\ell-\frac{\pi}{M^2}}^{\ell+\frac{\pi}{M^2}}\frac{\cos\left(\frac{tu}{M}\right)-1}{t}\mathcal{B}_1(t)\,dt \\ &+\frac{1}{\pi}\int_0^{\frac{\pi}{M^2}}\frac{\cos\left(\frac{tu}{M}\right)-1}{t}\mathcal{B}_1(t)\,dt \\ &+\frac{1}{\pi}\int_{(0,M\pi)/A}\frac{\cos\left(\frac{tu}{M}\right)-1}{t}\mathcal{B}_1(t)\,dt,\end{aligned} \quad (70)$$

where $A = (0,\pi M^{-2}) \cup \left(\cup_{\ell=1}^{M}\left(\ell - \frac{\pi}{M^2}, \ell + \frac{\pi}{M^2}\right)\right)$. Since $|\mathcal{B}_1(t)| \le 1/2$, the first term on the right of (70) is bounded in absolute value by

$$G\sum_{\ell=1}^{M}\int_{\ell-\frac{\pi}{M^2}}^{\ell+\frac{\pi}{M^2}} t^{-1}dt = G\sum_{\ell=1}^{M}\left(\log\left(\ell+\frac{\pi}{M^2}\right) - \log\left(\ell-\frac{\pi}{M^2}\right)\right) = O\left(M^{-2}\log M\right),$$

by Taylor expansion of $\log(x)$ where G is a generic finite constant. By Taylor expansion of $\cos(x)$ the second term on the right of (70) is bounded in absolute value by

$$G\frac{u^2}{M^2}\int_0^{\frac{\pi}{M^2}} t\,dt = O\left(M^{-4}\right).$$

On the other hand, since $\lim_q \sum_{v=1}^{q}\frac{\sin(2\pi vt)}{v} = -\pi\mathcal{B}_1(t)$ uniformly in $0 < t < 1$, see Courant and John (1974, p.635), the third term on the right of (70) is

$$\begin{aligned}&\frac{1}{\pi^2}\sum_{v=1}^{\infty}\frac{1}{v}\int_{(0,M\pi)/A}\frac{1-\cos\left(\frac{tu}{M}\right)}{t}\sin(2\pi vt)\,dt \\ =\ &\frac{1}{2\pi^2}\sum_{v=1}^{\infty}\frac{1}{v}\int_{(0,M\pi)} t^{-1}\left(2\sin(2\pi vt) - \sin\left(2\pi vt + \frac{tu}{M}\right) - \sin\left(2\pi vt - \frac{tu}{M}\right)\right)dt \\ &+O\left(M^{-2}\log M\right)\end{aligned} \quad (71)$$

using that $\cos(a)\sin(b) = 2^{-1}(\sin(a+b)+\sin(b-a))$ as we now show. First we examine that

$$\sum_{v=1}^{\infty}\frac{1}{v}\int_0^{\frac{\pi}{M^2}}\frac{1-\cos\left(\frac{tu}{M}\right)}{t}\sin(2\pi vt)\,dt = O\left(M^{-2}\right).$$

By integration by parts the left side of the last displayed equation is

$$-\sum_{v=1}^{\infty}\frac{1}{2\pi v^2}\frac{1-\cos\left(\frac{tu}{M}\right)}{t}\cos(2\pi vt)\Big|_0^{\frac{\pi}{M^2}}$$
$$+\sum_{v=1}^{\infty}\frac{1}{2\pi v^2}\int_0^{\frac{\pi}{M^2}}\cos(2\pi vt)\left(\frac{\frac{u}{M}\sin\left(\frac{tu}{M}\right)}{t}-\frac{1-\cos\left(\frac{tu}{M}\right)}{t^2}\right)dt.$$

The first term by Taylor expansion of $1-\cos(z)$ is bounded in absolute value by GM^{-2} since $0\le u\le M$, whereas by Taylor expansion of $\sin(z)$ and $1-\cos(z)$ around $z=0$ the second term is bounded in absolute value by

$$\sum_{v=1}^{\infty}\frac{G}{v^2}\int_0^{\frac{\pi}{M^2}}\left(\frac{u}{M}\right)^2 dt = O\left(M^{-2}\right).$$

Next, in the set $\cup_{\ell=1}^{M}\left(\ell-\frac{\pi}{M^2},\ell+\frac{\pi}{M^2}\right)$. Denoting

$$h_1(t)=\frac{1-\cos\left(\frac{ut}{M}\right)}{t}\quad\text{and}\quad h_2(t)=\cos(2\pi vt)$$

we need to show that

$$-\sum_{v=1}^{\infty}\frac{1}{2\pi v^2}\sum_{\ell=1}^{M}h_1(t)\,h_2(t)\Big|_{\ell-\frac{\pi}{M^2}}^{\ell+\frac{\pi}{M^2}}$$
$$+\sum_{v=1}^{\infty}\frac{1}{2\pi v^2}\sum_{\ell=1}^{M}\int_{\ell-\frac{\pi}{M^2}}^{\ell+\frac{\pi}{M^2}}\cos(2\pi vt)\left(\frac{\frac{u}{M}\sin\left(\frac{tu}{M}\right)}{t}-\frac{1-\cos\left(\frac{tu}{M}\right)}{t^2}\right)dt = O\left(\frac{1}{M}\right).$$

The second term is easily shown to be $O\left(M^{-1}\right)$, whereas since $h_2\left(\ell+\frac{\pi}{M^2}\right)=h_2\left(\ell-\frac{\pi}{M^2}\right)$ the first term is

$$-\sum_{v=1}^{\infty}\frac{1}{2\pi v^2}\sum_{\ell=1}^{M}\left(h_1\left(\ell+\frac{\pi}{M^2}\right)-h_1\left(\ell-\frac{\pi}{M^2}\right)\right)h_2\left(\ell-\frac{\pi}{M^2}\right)=O\left(M^{-1}\right)$$

by Taylor expansion of $h_1(z)$ around $z=\ell-\frac{\pi}{M^2}$.

So, we need to examine the first term on the right of (71). Since $\int_0^\infty t^{-1}\sin(t)\,dt = 2^{-1}\pi$ implies that

$$\begin{aligned} &2\int_{2\pi^2 Mv}^\infty t^{-1}\sin(t)\,dt - \int_{2\pi^2 Mv+u\pi}^\infty t^{-1}\sin(t)\,dt - \int_{2\pi^2 Mv-u\pi}^\infty t^{-1}\sin(t)\,dt \\ =& \int_{2\pi^2 Mv-u\pi}^\infty t^{-1}\sin(t)\,dt + \int_{2\pi^2 Mv+u\pi}^\infty t^{-1}\sin(t)\,dt - 2\int_{2\pi^2 Mv}^\infty t^{-1}\sin(t)\,dt, \end{aligned}$$

the first term on the right of (71) is bounded by

$$G\sum_{v=1}^\infty \frac{1}{v}\left(\left|\int_{2\pi^2 Mv+u\pi}^\infty t^{-1}\sin(t)\,dt\right| + \left|\int_{2\pi^2 Mv-u\pi}^\infty t^{-1}\sin(t)\,dt\right| + \left|\int_{2\pi^2 Mv}^\infty t^{-1}\sin(t)\,dt\right|\right) = O\left(M^{-1}\right)$$

because $\left|\int_a^\infty t^{-1}\sin(t)\,dt\right| \le Ga^{-1}$, see Courant and John (1974), and $0 \le u \le M$. This completes the proof of part (a).

Next part (b). Proceeding as with part (a), after an obvious change of variables, the left side of (69) is

$$\frac{1}{M\pi}\left\{\int_A + \int_{(0,M\pi)/A}\right\}\sin\left(\frac{tu}{M}\right)\log\left(\frac{t}{M}\right)\mathcal{B}_1(t)\,dt. \qquad (72)$$

From the definition of $\mathcal{B}_1(t)$, the first term of (72) is bounded by

$$G\log M\int_0^{\pi M^{-2}}\left|t-\frac{1}{2}\right|dt = O\left(M^{-2}\log M\right).$$

Next the second term of (72). Writing $\mathcal{B}_1(t)$ as was done in part (a), this term is

$$-\frac{1}{M\pi^2}\sum_{v=1}^\infty v^{-1}\int_{\pi M^{-2}}^{M\pi}\sin\left(\frac{tu}{M}\right)\sin(2\pi vt)\log\left(\frac{t}{M}\right)dt + O\left(M^{-2}\log M\right), \qquad (73)$$

since the difference between the integral in $(\pi M^{-2}, M\pi)$ and $(0, M\pi)/A$ is easily shown to be $O\left(M^{-2}\log M\right)$ proceeding similarly as in part (a). Moreover

$$\begin{aligned} &\frac{1}{M}\sum_{v=1}^\infty v^{-1}\left|\int_0^{\pi M^{-2}}\sin\left(\frac{tu}{M}\right)\sin(2\pi vt)\log\left(\frac{t}{M}\right)dt\right| \\ \le& \frac{G}{M}\sum_{v=1}^\infty v^{-2}\left|\sin\left(\frac{tu}{M}\right)\cos(2\pi vt)\log\left(\frac{t}{M}\right)\right|_0^{M^{-2}\pi} \\ &+\frac{1}{M}\int_0^{\pi M^{-2}}\cos(2\pi vt)\left(\frac{u}{M}\cos\left(\frac{tu}{M}\right)\log\left(\frac{t}{M}\right) - \frac{\sin\left(\frac{tu}{M}\right)}{t}\right)dt\Bigg| = O\left(M^{-3}\log M\right) \end{aligned}$$

by Taylor expansion of $\sin\left(\frac{tu}{M}\right)$ up to its second term since $|u| \leq M$. So the last displayed inequality and that $\int_0^{\pi M^{-2}} |\log(t) - \log(M)|\, dt = O\left(M^{-2}\log M\right)$ implies that (73) is

$$-\frac{1}{M\pi^2}\sum_{v=1}^{\infty} v^{-1}\int_0^{M\pi}\sin\left(\frac{tu}{M}\right)\sin(2\pi vt)\log\left(\frac{t}{M}\right)dt + O\left(M^{-2}\log M\right).$$

But using the equality $\sin(a)\sin(b) = 2^{-1}(\cos(a-b) - \cos(a+b))$ and integration by parts the last expression is

$$\begin{aligned}
& \frac{1}{2\pi}\sum_{v=1}^{\infty}\frac{1}{v}\left(\frac{\sin\left(t\left(\frac{u}{M}+2\pi v\right)\right)}{(u+2M\pi v)} - \frac{\sin\left(t\left(2\pi v - \frac{u}{M}\right)\right)}{(2M\pi v - u)}\right)\log\left(\frac{t}{M}\right)\Bigg|_0^{M\pi} \\
& -\frac{1}{2\pi}\sum_{v=1}^{\infty}\frac{1}{v}\int_0^{M\pi}\left(\frac{\sin\left(t\left(\frac{u}{M}+2\pi v\right)\right)}{(u+2M\pi v)\,t} - \frac{\sin\left(t\left(2\pi v - \frac{u}{M}\right)\right)}{(2M\pi v-u)\,t}\right)dt + O\left(M^{-2}\log M\right) \\
= \; & -g_2(u) - \frac{1}{2\pi^2}\sum_{v=1}^{\infty}v^{-1}\left\{\frac{1}{2M\pi v + u}\int_{2M\pi^2 v + u\pi}^{\infty} - \frac{1}{2M\pi v - u}\int_{2M\pi^2 v - u\pi}^{\infty}\right\}\frac{\sin z}{z}dz \\
& -\frac{1}{4\pi}\sum_{v=1}^{\infty}v^{-1}\left\{\frac{1}{2M\pi v+u} - \frac{1}{2M\pi v - u}\right\} + O\left(M^{-2}\log M\right) \qquad (74) \\
= \; & g_1(u) - g_2(u) + O\left(M^{-2}\log M\right)
\end{aligned}$$

using the change of variable $\left(2\pi v - \frac{u}{M}\right)t = z$ and $\left(\frac{u}{M} + 2\pi v\right)t = z$ and that $\int_0^{\infty} z^{-1}\sin(z)\,dz = \pi/2$ in the first equality of (74), whereas for the second equality we have employed that $\left|\int_a^{\infty} z^{-1}\sin(z)\,dz\right| = O\left(a^{-1}\right)$ and then that

$$\sum_{v=1}^{\infty} v^{-1}\frac{1}{(2M\pi v \pm u)^2} = O\left(M^{-2}\right)$$

which concludes the proof of the Lemma. □

Lemma 7 *$a_{u,n} - a_u = O\left(M^{-1}\log M\right)$.*

Proof. $a_{u,n} - a_u$ is

$$\begin{aligned}
& \frac{1}{2M}\sum_{j=-M+1}^{M-1}\left(A_{j,n} - A_j^*\right)e^{iu\lambda_{2mj}} + \frac{1}{2M}\sum_{j=-M+1}^{M-1}\left(A_j^* - A_j\right)e^{iu\lambda_{2mj}} \\
& + \left(\frac{1}{2M}\sum_{j=-M+1}^{M-1}A_j e^{iu\lambda_{2mj}} - \frac{1}{2\pi}\int_{-\pi}^{\pi}A(\lambda)\,e^{iu\lambda}d\lambda\right), \qquad (75)
\end{aligned}$$

where $A_j^* = \exp\left(-\sum_{s=1}^{M-1} c_s e^{-is\lambda_{2mj}}\right)$. The third term of (75) is $O\left(M^{-1}\right)$ by Brillinger (1981, p.15) since $|A(\lambda)|\cos(u\lambda)$ and $|A(\lambda)|\sin(u\lambda)$ have an integrable derivative.

Assume that $\alpha > 0$, the proof for $\alpha = 0$ is identical. The second term of (75) is

$$\frac{1}{2M}\sum_{j=-M+1}^{M-1} A_j\left(\exp\left\{\sum_{s=M}^{\infty} c_s e^{-is\lambda_{2mj}}\right\}-1\right)e^{iu\lambda_{2mj}}$$

which is bounded in absolute value by

$$\frac{G}{2M}\sum_{j=-M+1;j\neq 0}^{M-1}|A_j|\left|\sum_{s=M}^{\infty} c_s e^{-is\lambda_{2mj}}\right|(1+O(1)) \leq \frac{G}{2M^2}\sum_{j=-M+1;j\neq 0}^{M-1}\lambda_{2mj}^{-1}|A_j| = O\left(M^{-1}\right)$$

since $A(0)=0$, Lemma 5 implies that $\sum_{s=M}^{\infty} c_s e^{-is\lambda_{2mj}} = O\left(\lambda_{2mj}^{-1}M^{-1}\right)$ and $C.1$ that $|A_j| = O\left((j/M)^{\alpha/2}\right)$.

To complete the proof we need to examine the first term of (75) which by definition of $A_{j,n}$ and A_j^* is

$$-\frac{1}{2M}\sum_{j=-M+1}^{M-1}\left(1-\exp\left\{-\sum_{s=1}^{M-1}(c_{s,n}-c_s)\,e^{-is\lambda_{2mj}}\right\}\right)A_j^* e^{iu\lambda_{2mj}}. \qquad (76)$$

Now, from the definition of $c_{s,n}$ and c_s, we have that $\sum_{s=1}^{M-1}(c_{s,n}-c_s)\,e^{-is\widetilde{\lambda}_j}$ is

$$\begin{aligned}
&\sum_{s=1}^{M-1}\left(\frac{1}{M}\sum_{1\leq\ell<M}\log(f_{x,\ell})\cos\left(s\widetilde{\lambda}_\ell\right)-\frac{1}{\pi}\int_0^\pi \log f_x(\lambda)\cos(s\lambda)\,d\lambda\right)e^{-is\widetilde{\lambda}_j}\\
=&\sum_{s=1}^{M-1}\left(\frac{1}{M}\sum_{1\leq\ell<M}\log(g_{x,\ell})\cos\left(s\widetilde{\lambda}_\ell\right)-\int_0^1(\log g_x(\lambda\pi))\cos(s\lambda)\,d\lambda\right)e^{-is\widetilde{\lambda}_j}\\
&-\alpha\sum_{s=1}^{M-1}\left(\frac{1}{M}\sum_{1\leq\ell<M}\log\left(\widetilde{\lambda}_\ell\right)\cos\left(s\widetilde{\lambda}_\ell\right)-\int_0^1\log(\pi\lambda)\cos(s\lambda\pi)\,d\lambda\right)e^{-is\widetilde{\lambda}_j} \qquad (77)
\end{aligned}$$

by $C.1$. Since by $C.1$ $g(\lambda)$ is twice continuously differentiable the first term of (77) is

$$\begin{aligned}
&\frac{1}{M}\sum_{s=1}^{M-1}(\log g(0)+\cos(s\pi)\log g(\pi))\,e^{-is\widetilde{\lambda}_j}+O\left(M^{-1}\right)\\
=\;& M^{-1}\left(D\left(-\widetilde{\lambda}_j\right)\log g(0)+2^{-1}\left(D\left(\widetilde{\lambda}_j+\pi\right)+D\left(\widetilde{\lambda}_j-\pi\right)\right)\log g(\pi)\right)+O\left(M^{-1}\right)
\end{aligned}$$

by Brillinger (1981, p.15) where $D(\lambda)=\sum_{s=1}^{M-1}e^{is\lambda}$. Because $\left|\sum_{s=1}^{M-1}e^{-is\lambda_{2mj}}\right| \leq \lambda_{2mj}^{-1}$ and $\left|A_j^*\right| = O\left(\widetilde{\lambda}_j^{\alpha/2}\right)$, the contribution of the first term of (77) into (76) is bounded in absolute value by

$$\frac{G}{M}\sum_{j=1}^{M-1} j^{-1}\widetilde{\lambda}_j^{\alpha/2} = O\left(M^{-1}\right).$$

To complete the proof we need to examine the contribution in (76) due to the second term on the right of (77). That term is

$$-\alpha\sum_{s=1}^{M-1} e^{-is\widetilde{\lambda}_j}\left(\frac{1}{M}\sum_{1\le\ell<M}\log\left(\widetilde{\lambda}_\ell\right)\left(\cos\left(s\widetilde{\lambda}_\ell\right)-1\right)-\int_0^1\log\left(\lambda\pi\right)\left(\cos\left(s\lambda\pi\right)-1\right)d\lambda\right)$$
$$-\alpha\left(\frac{1}{M}\sum_{1\le\ell<M}\log\left(\widetilde{\lambda}_\ell\right)-\int_0^1\log\left(\lambda\pi\right)d\lambda\right)D\left(\widetilde{\lambda}_j\right). \tag{78}$$

The second term of (78) is $O\left(j^{-1}\log M^{\alpha/2}\right)$ since $\left|D\left(\widetilde{\lambda}_j\right)\right|\le\widetilde{\lambda}_j^{-1}$ and by Robinson's (1995) Lemma 2 $M^{-1}\sum_{1\le\ell<M}\log\left(\widetilde{\lambda}_\ell\right)-\int_0^1\log\left(\lambda\pi\right)d\lambda = O\left(M^{-1}\log M\right)$. Next, the first term of (78). Since the function $(\cos(u)-1)\log(u)$ has an integrable derivative, by Brillinger (1981, $p.$15), the first term of (78) is

$$-\frac{\alpha\log(\pi)}{2M}\sum_{s=1}^{M-1}\left(\cos\left(s\pi\right)-1\right)e^{-is\widetilde{\lambda}_j} \tag{79}$$
$$-\frac{\alpha}{M}\sum_{s=1}^{M-1}\int_0^1\mathcal{B}_1\left(\pi M\lambda\right)\frac{\cos\left(s\lambda\pi\right)-1}{\lambda}d\lambda e^{-is\widetilde{\lambda}_j}$$
$$+\frac{\alpha}{M}\sum_{s=1}^{M-1}s\pi\int_0^1\mathcal{B}_1\left(\pi M\lambda\right)\sin\left(s\lambda\pi\right)\log\left(\lambda\pi\right)d\lambda e^{-is\widetilde{\lambda}_j}.$$

Because $(1-\cos(s\pi)) = 2\mathcal{I}\left(s = odd\right)$, the first term of (79) is $O\left(\alpha j^{-1}\right) = O\left(j^{-1}\log M^{\alpha/2}\right)$, whereas by Lemma 6 part (a) the second term is $O\left(M^{-1}\right)$. Finally by Lemma 6 part (b) the third term of (79) is

$$\frac{\alpha\pi}{M}\sum_{s=1}^{M-1}s\left(M^{-2}\log M+\left(g_1(s)-g_2(s)\right)e^{-is\widetilde{\lambda}_j}\right) = O\left(\frac{\log M}{M}\right)+\frac{\alpha\pi}{M}\sum_{s=1}^{M-1}sg_1(s)e^{-is\widetilde{\lambda}_j}$$
$$-\frac{\alpha\pi}{M}\sum_{s=1}^{M-1}sg_2(s)e^{-is\widetilde{\lambda}_j}.$$

By Abel summation by parts the second term is bounded in absolute value by

$$\frac{\alpha}{j}\sum_{s=1}^{M-1}\left|sg_1(s)-(s+1)g_1(s+1)\right| = O\left(\frac{\alpha}{j}\sum_{s=1}^{M-1}\left(sM^{-2}+M^{-2}\right)\right) = O\left(\alpha j^{-1}\right)$$

as is easily shown from the definition of $g_1(s)$ in Lemma 6, whereas the third term is

$$-\frac{\alpha\pi}{M}\sum_{s=1}^{[M/2]}(2s-1)\,g_2\left((2s-1)\right)e^{-i(2s-1)\widetilde{\lambda}_j}-\frac{\alpha\pi}{M}\sum_{s=1}^{[M/2]}2sg_2\left(2s\right)e^{-i2s\widetilde{\lambda}_j}=O\left(\alpha j^{-1}\right)$$

proceeding as we did for $g_1(s)$ to each of the terms on the left, and where for notational simplicity we have assumed that M is even.

So, we conclude that the contribution of the second term of (77) into (76) is bounded in absolute value by

$$\begin{aligned}
& \frac{G}{M}\sum_{j=1}^{M}\left|1-\exp\left(j^{-1}\log M^{\alpha/2}+M^{-1}\log M\right)\right|\widetilde{\lambda}_j^{\alpha/2} \\
\leq\;& \frac{GM^{\alpha/2}}{M}\sum_{j=1}^{\log M}j^{-1}\widetilde{\lambda}_j^{\alpha/2}+\frac{G}{M}\sum_{j=1+\log M}^{M/\log M}j^{-1}\widetilde{\lambda}_j^{\alpha/2}+\frac{G\log M}{M^2}\sum_{j=1+M/\log M}^{M}\widetilde{\lambda}_j^{\alpha/2} \\
=\;& O\left(M^{-1}\log M\right)
\end{aligned}$$

since $|\alpha|<1$ and $|1-\exp(z)|\leq G|z|$ if z is bounded. □

8. CONCLUSIONS AND EXTENSIONS

In this paper we have extended the nonparametric prediction algorithm examined by Bhansali (1974, 77) to any covariance stationary linear process which may exhibit strong dependence. Since we do not impose any particular structure on the underline process of the data, we are thus able to avoid the problem that the misspecification of the model may induce to obtain adequate predictions. In addition, we have discussed how the $FLES$ can be adapted to extract the signal from a covariance stationary strong dependent process. One feature, in contrast to previous work on the topic, is that we do not need to assume any particular model for the noise. So, we can coin the approach as semiparametric.

An alternative method to predict x_t or recover the signal is via the estimation of the spectral density function by fitting an autoregressive $AR(P)$ model where P increases with the sample size, see Berk (1974) or An *et al.* (1982) among others. However, as Bhansali (1978) showed for weakly dependent data, this method is asymptotically equivalent to that described in Section 2, at least for prediction purposes. It thus seems to be of interest to examine whether the results hold the same under strong dependence.

REFERENCES

[1] An, H-Z., Z-G. Chen, and E.J. Hannan (1982) Autocorrelation, autoregression and autoregressive approximation. *Annals of Statist*ics 10, 926-936.

[2] Andel, J. (1986) Long-memory time series models. *Kybernetika* 22, 105-123.

[3] Arteche, J. and Robinson, P.M. (2000) Semiparametric inference in seasonal and cyclical long memory processes. *Journal of Time Series Analysis*, 21, 1-25.

[4] Beran, J. (1994) Statistics for long memory processes. New York: Chapman and Hall.

[5] Berk, K.N. (1974) Consistent autoregressive spectral estimates. *Annals of Statist*ics 2, 489-502.

[6] Bhansali, R.J. (1974) Asymptotic properties of the Wiener-Kolmogorov predictor. I. *Journal of the Royal Statistical So*ciety *Series* B, 36, 61-73.

[7] Bhansali, R.J. (1977) Asymptotic properties of the Wiener-Kolmogorov predictor. II. *Journal of the Royal Statistical Society Series* B, 39, 66-72.

[8] Bhansali, R.J. (1978) Linear prediction by autoregressive model fitting in the time domain. *Annals of Statist*ics 6, 224-231.

[9] Bloomfield, P. (1973) An exponential model for the spectrum of a scalar time series. *Biometrika* 60, 217-226.

[10] Brillinger, D.R. (1981) Time series, data analysis and theory. San Francisco: Holden Day.

[11] Carlin, J.B. and A.P. Dempster (1989) Sensitivity analysis of seasonal adjustments: Empirical case studies. *Journal of the American Statistical Association* 84, 6-20.

[12] Cleveland, W.P. and G.C. Tiao (1976) Decomposition of seasonal time series: A model for the Census X-11 program. *Journal of the American Statistical Association*, 71, 581-587.

[13] Courant, R. and F. John (1974) Introducción al Análisis Matemático. Vol. 1. Limusa, México.

[14] Crato, N. and B.K. Ray (1996) Model selection and forecasting for long-range dependent processes. *Journal of Forecasting* 15, 107-125.

[15] Dahlhaus, R. (1989) Efficient parameter estimation for self-similar processes. *Annals of Statist*ics 17, 1749-1766.

[16] Fox, R. and M.S. Taqqu (1986) Large sample properties of parameter estimates for strongly dependent stationary Gaussian time series. *Annals of Statist*ics 14, 517-532.

[17] Giraitis, L., Hidalgo, J. and P.M. Robinson (2000) Gaussian estimation of parametric spectral density with unknown pole. Preprint.

[18] Giraitis, L. and D. Surgailis (1990) A central limit theorem for quadratic forms in strongly dependent linear variables and its application to asymptotic normality of Whittle's estimate. *Probability Theory and Related Fields* 86, 87-104.

[19] Gray, H.L., N. Zhang and W.A. Woodward (1989) On generalized fractional processes. *Journal of Time Series Analysis* 10, 233-257.

[20] Hannan, E.J. (1970) Multiple time series. Wiley, New York.

[21] Hidalgo, J. (2000a) Nonparametric test for causality with long range dependence. *Econometrica* 68, 1465-1490..

[22] Hidalgo, J. (2000b) Semiparametric estimation of the location of the pole. Preprint.

[23] Hidalgo, J. and Y. Yajima (1999) Semiparametric estimation of the long-range parameter. Preprint.

[24] Hosoya, Y. (1997) A limit theory with long-range dependence and statistical inference on related models. *Annals of Statist*ics 25, 105-137.

[25] Mandelbrot, B.B. and J.W. Van Ness (1968) Fractional Brownian motions, fractional noises and applications. *SIAM Review* 10, 422-437.

[26] Peiris, M.S. and B.J.C. Perera (1988) On prediction with fractionally differenced ARIMA models. *Journal of Time Series Analysis* 9, 215-220.

[27] Porter-Hudak, S. (1990) An application of the seasonal fractionally differenced model to the monetary aggregates. *Journal of the American Statistical Association* 85, 338-344.

[28] Press, H. and J.W. Tukey (1956) Power Spectral methods of analysis and their application to problems in airplane dynamics. Flight Test Manual, NATO, Advisory Group for Aeronautical Research and Development, Vol. IV-C, 1-41. Reprinted as Bell System Monograph No. 2606.

[29] Ray, B.K. (1993) Long range forecasting of IBM product revenues using a seasonal fractional differenced ARMA model. *International Journal of Forecasting* 9, 255-269.

[30] Robinson, P.M. (1994) Time series with strong dependence. In C.A. Sims, ed., Advances in Econometrics: Sixth World Congress, Vol. 1, 47-95. Cambridge University Press, Cambridge.

[31] Robinson, P.M. (1995) Gaussian semiparametric estimation of long range dependence. *Annals of Statist*ics 23, 1630-1661.

[32] Serfling, R.J. (1980) Approximation theorems of mathematical statistics. John Wiley and Sons. New York.

[33] Sinai, Y.G. (1976) Self-similar probability distribution. *Theory of Probability and its Applications* 21, 64-80.

[34] Taniguchi, M. (1987) Minimum contrast estimation for spectral densities of stationary processes. *Journal of the Royal Statistical Society Series* B, 49, 315–325.

[35] Whittle, P. (1963) Prediction and regulation by linear least squares methods. Princeton: Van Nostram.

[36] Yajima, Y. (1996) Estimation of the frequency of unbounded spectral densities. Proceedings of the Business and Economic Statistical Section. *American Statistical Association*.